THE QUARTERLY JOURNAL OF MATHEMATICS

OXFORD SECOND SERIES

Volume 1 *No.* 2 *June* 1950

OXFORD

AT THE CLARENDON PRESS

Price 10*s.* 6*d.* *net*

PRINTED IN GREAT BRITAIN BY CHARLES BATEY AT THE OXFORD UNIVERSITY PRESS

THE QUARTERLY JOURNAL OF MATHEMATICS

OXFORD SECOND SERIES

Edited by T. W. CHAUNDY, U. S. HASLAM-JONES, J. H. C. THOMPSON

THE QUARTERLY JOURNAL OF MATHEMATICS (OXFORD SECOND SERIES) is published at 10s. 6d. net for a single number with an annual subscription (for four numbers) of 37s. 6d. post free.

Papers, of a length normally not exceeding 20 printed pages of the Journal, are invited on subjects of Pure and Applied Mathematics, and should be addressed 'The Editors, Quarterly Journal of Mathematics, Clarendon Press, Oxford'. The Editors as a rule will not wish to accept material that they cannot see their way to publish within a twelvemonth. While every care is taken of manuscripts submitted for publication, the Publisher and the Editors cannot hold themselves responsible for any loss or damage. Authors are advised to retain a copy of anything they may send for publication. Authors of papers printed in the Quarterly Journal will be entitled to 50 free offprints. Correspondence on the *subject-matter* of the Quarterly Journal should be addressed, as above, to 'The Editors', at the Clarendon Press. All other correspondence should be addressed to the Publisher

GEOFFREY CUMBERLEGE

OXFORD UNIVERSITY PRESS

AMEN HOUSE, LONDON, E.C. 4

A GEOMETRICAL PROPERTY OF SETS OF FRACTIONAL DIMENSION

By H. G. EGGLESTON (*Swansea*)

[Received 3 February 1949]

Introduction

It is a well-known property of Lebesgue measure that, if, in a Euclidean plane with rectangular coordinate axes x and y, a measurable set A is given on the x-axis and a measurable set B on the y-axis, both of finite linear measure, then the planar measure of the Cartesian product set is equal to the product of the linear measures of A and B. Results of this type have been extended to other measures, as in the references (1)–(5).

In this paper a result is established about a related problem which may be stated as follows.

If A is a given set of finite positive s-measure and through each point of A an arc of length l is drawn, what are the dimension and measure of the set of points which constitute all the arcs?

The definition of s-measure is as follows. For a point set A and $\delta > 0$, let $U(\delta, A)$ denote a collection of point sets whose point-set sum contains A and each member of which has diameter less than δ. $I(\delta, A)$ denotes the class of all $U(\delta, A)$. Then

$$\Lambda_s^* = \lim_{\delta\to 0}\Big\{\underset{I(\delta,A)}{\text{lower bound}} \sum_{U(\delta,A)} d^s\Big\}$$

where d denotes the diameter of a typical member of $U(\delta, A)$ and the summation is over all these members.

If A is Λ_s-measurable, we write Λ_s for Λ_s^*. A set which is planar, measurable Λ_s, and of finite positive s-measure is called an 's-set'.

A $U(\delta, A)$ whose members are convex is denoted by $\chi(\delta, A)$; $\rho(x, y)$ is used for the distance between two points x and y; $\lambda(x, y)$ for the arc length between the points x and y (used only when it is clear from the context which arc is under consideration). The diameter of a point set A is written $d(A)$ and the complement of A in the plane is written $b(A)$.

1. Theorem. *Suppose that*

(i) *A is an s-set*;

(ii) *each point p of A is the end-point of an arc $L(p)$ of length l, which also lies in the plane*;

(iii) *k is a constant* $(0 < k \leqslant 1)$;

(iv) $\rho(x_1, x_2) \geqslant k\rho(p_1, p_2)$

for all p_1, p_2 of A, and all x_1 of $L(p_1)$, all x_2 of $L(p_2)$.

Then
$$L(A) = \sum_{p \in A} L(p)$$
is a set of dimension at least $s+1$ and, if it is of dimension $s+1$, then
$$\Lambda^*_{s+1}\{L(A)\} \geqslant k^s l \Lambda_s(A). \tag{1}$$

The method is to show that, for any $\epsilon > 0$,
$$\Lambda^*_{s+1}\{L(A)\} \geqslant k^s l \Lambda_s(A)(1-\epsilon). \tag{2}$$
This will be so if, given $\epsilon > 0$, there is a $\delta > 0$ such that for any $\chi\{\delta, L(A)\}$,
$$\sum_{\chi\{\delta, L(A)\}} d^{1+s} \geqslant k^s l \Lambda_s(A)(1-\epsilon). \tag{3}$$

Define ϵ_1 to be such that
$$0 < 3\epsilon_1 < 1-(1-\epsilon)^{\frac{1}{2}}. \tag{4}$$

2. Lemma 1. *Given an s-set A and $\epsilon_1 > 0$, there is a positive number δ_1 and a sub-set A_1 of A such that $\Lambda_s(A_1) > \Lambda_s(A)(1-\epsilon_1)$ and, for every convex set V of diameter less than or equal to δ_1,*
$$\{d(V)\}^s > (1-\epsilon_1)\Lambda_s(A_1 V). \tag{5}$$
For a proof see (**1**).

Lemma 2. *If A is the set of the theorem, and A_1, ϵ_1 are as in Lemma 1, then there is a $\delta_2 > 0$ and a sub-set A_2 of A_1 with the following properties.*

(i) $\Lambda^*_s(A_2) > \Lambda_s(A_1)(1-\epsilon_1)$,

(ii) *if p is a point of A_2, and $P(p)$ is a δ_2-polygonal line inscribed in $L(p)$, then $P(p)$ is of length greater than $l(1-\epsilon)$.*

By a δ_2-polygonal line is meant here a polygonal line whose sides are of length less than δ_2. The proof is omitted.

Let
$$\delta k^{-1} = \min\{\delta_1, \delta_2, l\epsilon_1\} \tag{6}$$
where δ_1, ϵ_1, and δ_2 are as in Lemmas 1, 2. It will be shown that (3) holds with respect to this δ.

3. Lemma 3. *If A_2 and ϵ_1 are as in Lemma 2 and $\chi\{\delta, L(A)\}$ is given, there is a sub-set A_3 of A_2 and a finite sub-set χ_1 of $\chi\{\delta, L(A)\}$ with the following properties.*

(i) $\Lambda^*_s(A_3) > \Lambda^*_s(A_2)(1-\epsilon_1)$;

(ii) *if, for p belonging to A_3, χ^p_1 is the sub-set of those sets of χ_1 which intersect $L(p)$, then*
$$\sum_{\chi_1^p} d > l(1-3\epsilon_1).$$

Proof. It may be supposed that $\chi\{\delta, L(A)\}$ consists of at most an enumerable number of sets, say $V_1, V_2,\ldots$. For each p of A the point-set intersection

$$L(p)\Big(\sum_{i=1}^{N} V_i\Big)$$

is increasing and tends to $L(p)$ as N tends to infinity. The integer N can be chosen so large that

$$\Lambda_s^*\{A(N)A_2\} > (1-\epsilon_1)\Lambda_s^*(A_2) \tag{7}$$

where $A(N)$ is the sub-set of those p of A for which

$$\Lambda\Big\{L(p)\Big(\sum_{i=1}^{N} V_i\Big)\Big\} > l(1-\epsilon_1). \tag{8}$$

The sets χ_1 and A_3 are taken to be $\{V_i\}$ ($i = 1,\ldots, N$) and $A(N)A_2$ respectively. It will be shown that the lemma is true with this χ_1 and A_3.

For any p of A let the members of χ_1^p be enumerated as $T_1, T_2,\ldots, T_{n_p}$. Let the boundary of T_i be called K_i. The points of $L(p)$ may be ordered with p as the first point. 'Before', 'after', and 'last' refer to this order. Denote by p' the last point of $L(p)$.

The vertices of a polygonal line $P(p)$ are defined as follows: y_0 is p, and, when y_r has been defined, y_{r+1} is defined (or not (α)) by the first of the five processes (β),..., (κ) (unless it is p') which is applicable:

(α) if y_r is p', the definition of the vertices is complete;

(β) if y_r is such that $0 < \rho(y_r, p') \leqslant \delta$, then y_{r+1} is p';

(γ) let D_r be $\sum K_i$ summed over those integers i for which T_i+K_i contains y_r; if D_r intersects $L(p)$ in a point after y_r, then y_{r+1} is the last of such points;

(ϕ) if the next point of intersection of $\sum_{i=1}^{n_p} K_i$ with $L(p)$ after y_r is z and if $\rho(y_r, z) \leqslant \delta$, then y_{r+1} is z;

(θ) if the next point of intersection of $\sum_{i=1}^{n_p} K_i$ with $L(p)$ after y_r is z and if $\rho(y_r, z) > \delta$, then y_{r+1} is the last point of $L(p)$ which comes before z and is such that $\rho(y_r, y_{r+1}) = \delta$;

(κ) if $\sum_{i=1}^{n_p} K_i$ has no points of intersection with $L(p)$ after y_r and if $\rho(y_r, p') > \delta$, then y_{r+1} is the last point of $L(p)$ for which $\rho(y_r, y_{r+1}) = \delta$.

There are at most a finite number of points y_r; let them be $y_0, y_1,\ldots, y_{r_p}$. The polygonal line $P(p)$ is obtained by joining y_r to y_{r+1} by a straight segment for $r = 0, 1, 2,\ldots, r_p-1$.

$P(p)$ has the following properties:

(i) each segment of $P(p)$ has length not more than δ and thus not more than δ_2, by (6);

(ii) if y_{r+1} is defined by (ϕ), (θ), or (κ), then the arc $y_r y_{r+1}$ is exterior to χ_1^p;

(iii) if y_{r+1} is defined by (γ), the segment $y_r y_{r+1}$ is completely contained in one of the T_i and no other segment of $P(p)$ is contained in the same T_i.

Denote by I the set of integers $0, 1,..., r_p-1$; by I' the sub-set of those r of I for which y_{r+1} is defined by (γ); by I'' the sub-set of those r of I for which y_{r+1} is defined by (ϕ), (θ), or (κ).

By (ii) above and (7),

$$\sum_{I''} \rho(y_r, y_{r+1}) \leqslant \sum_{I''} \lambda(y_r, y_{r+1}) \leqslant l\epsilon_1. \tag{9}$$

By (i) above and Lemma 2

$$\sum_{I} \rho(y_r, y_{r+1}) > l(1-\epsilon_1).$$

Thus

$$\sum_{I'} \rho(y_r, y_{r+1}) > l(1-2\epsilon_1)-\delta. \tag{10}$$

But, by (iii),

$$\sum_{\chi_1^p} d \geqslant \sum_{I'} \rho(y_r, y_{r+1}).$$

Hence, by (6),

$$\sum_{\chi_1^p} d > l(1-2\epsilon_1)-\delta \geqslant l(1-3\epsilon_1). \tag{11}$$

4. Proof of the theorem

Let $\chi = \chi\{\delta, L(A)\}$ be a covering of $L(A)$ by convex sets and let χ_1 and A_3 be defined so that Lemma 3 applies to them. Write the areas of χ_1 as $V_1, V_2,..., V_N$. Let the set of those p of A for which $L(p)$ meets V_j be called D_j, and denote W_j the smallest closed convex set containing D_j. Then

$$d(V_j) \geqslant k\, d(W_j).$$

The symbol $\gamma(i_1, i_2,..., i_N)$ denotes a point set defined for $i_j = 0, 1$ as follows:

$$\gamma(i_1, i_2,..., i_N) = \prod_{j=1}^{N} X_j$$

where, if $i_j = 1$, X_j is W_j, and, if $i_j = 0$, X_j is bW_j.

Then

$$W_j = \sum_{i_1=0}^{1} \cdots \sum_{i_{j-1}=0}^{1} \sum_{i_{j+1}=0}^{1} \cdots \sum_{i_N=0}^{1} \gamma(i_1,..., i_{j-1}, 1, i_{j+1},..., i_N),$$

which is written shortly as

$$W_j = \sum_{i_j=1} \gamma(i_1, i_2,..., i_N) \quad (j = 1, 2,..., N).$$

Thus
$$\begin{aligned}\{d(V_j)\}^s &\geqslant k^s\{d(W_j)\}^s\\ &\geqslant k^s(1-\epsilon_1)\Lambda_s(A_1 W_j)\\ &= k^s(1-\epsilon_1)\sum_{i_j=1}\Lambda_s\{A_1\gamma(i_1,\dots,i_N)\}\\ &\geqslant k^s(1-\epsilon_1)\sum_{i_j=1}\Lambda_s^*\{A_3\gamma(i_1,\dots,i_N)\}.\end{aligned}$$
Hence
$$\begin{aligned}\sum_{\chi\{\delta,\overline{L}(A)\}} d^{1+s} &\geqslant \sum_{j=1}^{N}\{d(V_j)\}^{1+s}\\ &\geqslant k^s(1-\epsilon_1)\sum_{j=1}^{N}\sum_{i_j=1}(dV_j)\Lambda_s^*\{A_3\gamma(i_1,\dots,i_N)\}\\ &\geqslant k^s(1-\epsilon_1)\sum{}'\,\Lambda_s^*\{A_3\gamma(i_1,\dots,i_N)\}\sum{}''\,d(V_j),\end{aligned}$$
where $\sum'$ is the sum over all possible sets of N terms $i_1, i_2,\dots, i_N$ for which $i_j = 0$ or 1 and at least one i_j is 1; $\sum''$ is the sum over all the integers j $(1 \leqslant j \leqslant N)$ for which $i_j = 1$ in the particular $\gamma(i_1,\dots, i_N)$ concerned.

If $A_3\gamma(i_1,\dots, i_N)$ is not void, let p be a point of it. $\sum'' d(V_j)$ is precisely the $\sum_{\chi_1^p} d$ of Lemma 3. Then $\sum'' dV_j > l(1-3\epsilon_1)$.

Hence
$$\begin{aligned}\sum_{\chi\{\delta,\overline{L}(A)\}} d^{1+s} &\geqslant k^s l(1-\epsilon_1)(1-3\epsilon_1)\sum{}'\,\Lambda_s^*\{A_3\gamma(i_1,\dots,i_N)\}\\ &\geqslant k^s l(1-\epsilon_1)(1-3\epsilon_1)\Lambda_s^*(A_3)\\ &\geqslant k^s l(1-\epsilon_1)^4(1-3\epsilon_1)\Lambda_s(A)\\ &\geqslant k^s l(1-\epsilon)\Lambda_s(A).\end{aligned}$$

This establishes the theorem.

Remark. The arcs in the theorem may be replaced by regular linear sets.

REFERENCES

1. A. S. Besicovitch and P. A. P. Moran, 'The measure of product and cylinder sets', *J. of London Math. Soc.* 20 (1945), 110–20.
2. P. A. P. Moran, 'Additive functions of intervals and Hausdorff measure', *Proc. Cambridge Phil. Soc.* 42 (1946), 15–23.
3. A. P. Morse and J. F. Randolph, 'Gillespie linear measure', *Duke Math. J.* 6 (1940), 408–19.
4. H. Federer, 'The (ϕ, k) rectifiable sub-sets of n-space', *Trans. American Math. Soc.* 62 (1947), 114–89.
5. P. A. P. Moran, 'On plane sets of fractional dimensions', *Proc. London Math. Soc.* 51 (1949), 415–23.

ON THE SPACE OF INTEGRAL FUNCTIONS (II)†

By V. GANAPATHY IYER (*Annamalai University, S. India*)

[Received 25 February 1949]

1. Introduction

Let E be a normed linear space.‡ Let S be a (linear) sub-space of E. Let $f(\alpha)$ $(\alpha \in S)$ be a continuous linear functional§ defined on S. Then there exists a continuous linear functional defined on the whole space E coinciding with $f(\alpha)$ on S [(B), 55, Theorem 2]. Also, if $f(\alpha) = 0$ for $\alpha \in S$ and $\alpha_0 \in E$ is at a positive distance from S, then there exists a continuous linear functional $F(\alpha)$ defined on E such that $F(\alpha_0) = 1$, $F(\alpha) = 0$, $\alpha \in S$ [(B), 57, Lemma]. The object of this paper is to extend these two results to the space of integral functions studied in (I) and to discuss some allied topics.‖ I first recall the definitions and notations of (I). I denote by Γ the class of all integral functions topologized by the distance $|\alpha-\beta|$ $(\alpha, \beta \in \Gamma)$ where

$$\alpha = \alpha(z) = \sum_0^\infty a_n z^n, \tag{1}$$

and $|\alpha|$ is defined†† by

$$|\alpha| = \text{upper bound}\,[|a_0|, |a_n|^{1/n}\ (n \geqslant 1)]. \tag{2}$$

With this topology, Γ is a non-normable complete linear metric space, and convergence in Γ is equivalent to uniform convergence in any finite circle [(I), Theorems 1, 2, 3]. I denote‡‡ by Γ^* the space of all continuous linear functionals defined on Γ. Every element $f \in \Gamma^*$ is of

† The first paper is printed in the *J. Indian Math. Soc.*, New Series, 12 (1948), 13–30. This will be referred to as (I) throughout this paper.

‡ For the definition of normed linear space see S. Banach, *Théorie des Opérations Linéaires*, Warsaw (1932), p. 53. I shall refer to this book as (B).

§ What I mean by 'continuous linear functional' is the 'linear functional' in (B). When I say *linear* I mean merely *algebraic linearity* with complex scalars.

‖ The two results stated above are proved in (B) for real linear space. They have been proved for complex linear spaces by Bohnenblust and Sobczyk; see *Bull. American Math. Soc.* 44 (1938), 91–3.

†† The symbol | | is used, in this paper as in (I), to denote (2) as well as the modulus of complex numbers. It will be evident from the context in which sense it is used.

‡‡ In (I) the adjoint was denoted by $\bar{\Gamma}$. Since we require the bar to denote closure, I have used Γ^* to denote the adjoint; but for this change the notation in (I) is strictly adhered to. In (I) a topology for Γ^* was introduced which made Γ an isometric sub-space of Γ^*. In this paper I do not consider any topology for Γ^*.

the form

$$f(\alpha) = \sum_0^\infty c_n a_n, \qquad \alpha = \sum_0^\infty a_n z^n, \tag{3}$$

where $\{|c_n|^{1/n}\}$ is bounded: (4)

that is, $f = f(z) = \sum_0^\infty c_n z^n$ has *positive* radius of convergence [(I), Theorem 4].

1.1 *Further definitions and notations.* Given an integral function

$$\alpha = \alpha(z) = \sum_0^\infty a_n z^n,$$

I define, for each $R > 0$, the expression $|\alpha; R|$ by the equation

$$|\alpha; R| = \sum_0^\infty |a_n| R^n. \tag{5}$$

It is easily seen that, for each $R > 0$, (5) defines a norm on the class of integral functions; I shall denote by $\Gamma(R)$ the normed linear space thus obtained and by $\Gamma^*(R)$ the adjoint of $\Gamma(R)$. If $R_1 > R_2$, then $|\alpha; R_1| \geqslant |\alpha; R_2|$, and so $\Gamma(R_1)$ is weaker† than $\Gamma(R_2)$ and

$$\Gamma^*(R_2) \subset \Gamma^*(R_1).$$

Moreover, if (α_p) be a sequence of integral functions such that $|\alpha_p| \to 0$ as $p \to \infty$, then, for each $R > 0$, $|\alpha_p; R| \to 0$ as $p \to \infty$. So $\Gamma(R)$ is stronger than Γ. Thus $\{\Gamma(R)\}$, as R increases, form a decreasing family of normed topologies on the class of integral functions each of which is stronger than Γ. In Theorem 1, I shall characterize Γ in terms of the family $\{\Gamma(R)\}$ and use it to prove the results stated in § 1.

2. If S be any class of integral functions and T any topology on the space of all integral functions, I shall denote by $(\bar{S})_T$ the closure of S in the topology T. When I speak of a 'sub-space of Γ' I mean a (complex) linear sub-space.

THEOREM 1. *Let S be any class of integral functions. Then*‡

$$(\bar{S})_\Gamma = \prod_{R>0} (\bar{S})_{\Gamma(R)}. \tag{6}$$

Proof. Since each $\Gamma(R)$ is stronger than Γ, it follows§ that

$$(\bar{S})_\Gamma \subset (\bar{S})_{\Gamma(R)}.$$

Hence $$(\bar{S})_\Gamma \subset B = \prod_{R>0} (\bar{S})_{\Gamma(R)}.$$

† In the sense of Alexandroff and Hopf: *Topologie*, I, Berlin (1935), 62. Cf. also R. Vaidyanathaswami, *Treatise on Set Topology*, part I, 71.

‡ The relation (6) shows that Γ is the lattice product [see R. Vaidyanathaswami, loc. cit. 130–5] of the family $\{\Gamma(R)\}$ of stronger topologies. But (6) is stronger than what is implied by Γ being the lattice-product of the family $\{\Gamma(R)\}$.

§ From the definition of stronger and weaker topologies.

To prove (6) we have therefore to prove that, if α is at a positive distance from S in Γ, then it is also at a positive distance from S in $\Gamma(R)$ *for some* $R > 0$ (and therefore also for all sufficiently large R). This is an immediate consequence of the following lemma, the proof of which will complete that of Theorem 1.

2.1. Lemma 1. *If* $|\alpha| \geqslant d > 0$, *then* $|\alpha; R| \geqslant d$ *for all* $R > A(1/d)$ *where* $A(t)$ *is defined, for* $t > 0$, *as* $\max(1, t)$.

Proof. Let $\alpha = \sum_0^\infty a_n z^n$ and $|\alpha| \geqslant d > 0$. If, for some $R > 0$, $|\alpha; R| < d$, it follows from (5) that

$$|a_0| < d, \qquad |a_n|^{1/n} R < d^{1/n} \leqslant A(d) \quad (n \geqslant 1). \tag{7}$$

If $A(d)/R < d$, we see from (7) that $|\alpha| < d$, contrary to hypothesis. Hence $A(d)/R \geqslant d$, that is,

$$R \leqslant A(d)/d = A(1/d).$$

Hence, if $R > A(1/d)$, we shall have $|\alpha; R| \geqslant d$, as is to be proved.

3. Theorem 2. $$\Gamma^* = \sum_{R>0} \Gamma^*(R).$$

3.1. To prove this theorem we require the following lemma.

Lemma 2. *Every functional in* $\Gamma^*(R)$ *is of the form*

$$f(\alpha) = \sum_0^\infty c_n a_n, \qquad \alpha = \sum_0^\infty a_n z^n \tag{8}$$

where $$\left\{\frac{|c_n|}{R^n}\right\} \text{ is bounded,} \tag{9}$$

and conversely.

Proof. Suppose that $f(\alpha)$ is a continuous linear functional on $\Gamma(R)$. Then there is a k such that $|f(\alpha)| \leqslant k|\alpha; R|$ [(B) 54]. Let $\delta_n = z^n$ and $f(\delta_n) = c_n$ $(n \geqslant 0)$. Then

$$f(\alpha) = \lim_{n\to\infty} (c_0 a_0 + \ldots + c_n a_n) = \sum_0^\infty c_n a_n,$$

since, in $\Gamma(R)$, $a_0 \delta_0 + \ldots + a_n \delta_n \to \alpha$ as $n \to \infty$. Also,

$$|c_n| \leqslant k|\delta_n; R| = kR^n.$$

Hence (9) is true. Conversely, if (9) is true, the functional defined by (8) exists for all α, and $|f(\alpha)| \leqslant k|\alpha; R|$ for some $k > 0$. Hence [(B), 54] $f(\alpha)$, which is obviously linear, is continuous on $\Gamma(R)$. So Lemma 2 is proved.

3.2. *Proof of Theorem* 2. Since a functional continuous in any topology will also be continuous in a weaker topology, it follows that

$\Gamma^*(R)$ belongs to Γ^* for each $R > 0$. Hence $\sum\limits_{R>0} \Gamma^*(R) \subset \Gamma^*$. To prove the reverse inclusion, let $f \in \Gamma^*$ be given by $f = \sum\limits_0^\infty c_n z^n$. Then, by (4), there is an $R > 0$ such that $|c_n|^{1/n} \leqslant R$, $n \geqslant 1$. Hence $\{|c_n|/R^n\}$ is bounded and $f \in \Gamma^*(R)$ by Lemma 2. So $\Gamma^* \subset \sum\limits_{R>0} \Gamma^*(R)$. This proves Theorem 2.

3.3. *Strong and weak convergence in* Γ. We have proved in (I) Theorem 8, that strong and weak convergence in Γ are equivalent. Theorem 2 leads to an alternative proof of this result. Let l denote the normed space of absolutely convergent series. If $\alpha = \sum\limits_0^\infty a_n z^n \in \Gamma(R)$ ($R > 0$ given), the transformation $a'_n = a_n R^n$ ($n \geqslant 0$) transforms $\Gamma(R)$ into an equivalent sub-space of l [see (B), 180 for the definition of equivalence]; let α' correspond to α. If $g \in l^*$, the functional $f(\alpha) = g(\alpha')$ belongs to $\Gamma^*(R)$. If, now, a sequence (α_p) of Γ converges to zero weakly, we have, by definition,† $f(\alpha_p) = g(\alpha'_p) \to 0$ for every $g \in l^*$. So, by a known result [(B), 137], $|\alpha'_p|_l \to 0$ as $p \to \infty$ where $|\alpha'|_l$ is the norm of α' in l. But $|\alpha_p; R| = |\alpha'_p|_l$. So $|\alpha_p; R| \to 0$. By Theorem 2, this is true for each R. Hence, by Lemma 1, $|\alpha_p| \to 0$ as $p \to \infty$. This proves the statement that strong and weak convergence are equivalent in Γ.

4. We can now prove the two theorems on the extension of functionals envisaged in § 1.

THEOREM 3. *Let S be a sub-space of* Γ. *Let $\alpha_0 \in \Gamma$ be at a distance $d > 0$ from S in* Γ. *Then for each $R > A(1/d)$ there is a functional $f \in \Gamma^*$ such that* (i) $f(\alpha_0) = 1$; (ii) $f(\alpha) = 0$ *for $\alpha \in S$; and* (iii) $|f(\alpha)| \leqslant |\alpha; R|/d$ *for all $\alpha \in \Gamma$, that is, $f \in \Gamma^*(R)$ for $R > A(1/d)$.*

Proof. Let $\beta \in S$. Then, by hypothesis, $|\alpha_0 - \beta| \geqslant d$. So, by Lemma 1 $|\alpha_0 - \beta; R| \geqslant d$ for $R > A(1/d)$. Hence the distance of α_0 from S in $\Gamma(R)$ exceeds or equals d for each $R > A(1/d)$. Using the known result for normed linear spaces [(B), 57, lemma, and above footnote ||, p. 86] we see that there is a functional $f \in \Gamma^*(R) \subset \Gamma^*$ satisfying the conditions stated in the theorem.

4.1. THEOREM 4. *Let S be a sub-space of* Γ. *Let $f(\alpha)$ be a linear functional defined and continuous (in the topology of* Γ) *on S. Then there is a functional $F \in \Gamma^*$ such that $F(\alpha) = f(\alpha)$ for $\alpha \in S$.*

† See (I) 22, § 7, for the definition of strong and weak convergence.

Proof. Since we know the truth of the corresponding result for normed spaces [(B), 55, Theorem 2], it is enough, by Theorem 2, to show that, if $f(\alpha)$ satisfies the conditions of the theorem, then it will be continuous on S in the topology of $\Gamma(R)$ for *some* $R > 0$. This we prove in the following lemma.

4.2. LEMMA 3. *Let S be a sub-space of Γ. Let $f(\alpha)$ be a linear functional defined and continuous on S in the topology of Γ. Then $f(\alpha)$ will be continuous on S in the topology of $\Gamma(R)$ for some $R > 0$.*

Proof. Suppose that $f(\alpha)$ is not continuous on S regarded as a sub-space of $\Gamma(R)$ for any $R > 0$. Then by known results for normed spaces [(B), 54, Theorem 1], we can, for each positive integer p, find an element $\alpha_p \in S$ such that (i) $|\alpha_p;p| = 1/p$; and (ii) $|f(\alpha_p)| \geqslant 1$. Arguing as in Lemma 1 we easily see that $|\alpha_p| \leqslant 1/p$. Hence $|\alpha_p| \to 0$, while $|f(\alpha_p)| \geqslant 1$ for all p, so that $f(\alpha)$ is not continuous on S contrary to hypothesis. This proves Lemma 3.

4.3. *Illustrations.* (i) Since it is easily proved that any finite set of linearly independent elements in Γ generate a closed linear sub-space, it follows by repeated application of Theorem 3 to one-dimensional sub-spaces that, given any linearly independent set $(\alpha_1,\ldots,\alpha_p)$ in Γ and arbitrary constants $(t_1, t_2,\ldots, t_p)$ there is a functional $f \in \Gamma^*$ such that $f(\alpha_i) = t_i$ $(i = 1, 2,\ldots, p)$. Of course, this can be proved directly without recourse to Theorem 3.

(ii) Let S denote the class of integral functions each of which vanishes at a given set of complex numbers having no finite limit point. By (I) Theorem 3, it follows that S is a proper closed linear sub-space of Γ. If α_0 is an integral function not belonging to S, then, by Theorem 3, there is an $f \in \Gamma^*$ such that $f(\alpha_0) = 1$, $f(\alpha) = 0$ for $\alpha \in S$. Explicitly, this asserts the existence of a solution (c_n) satisfying (4) of a system of infinite number of equations.

5. Bi-orthogonal sequences

A sequence (α_n) of Γ and a sequence (f_n) of Γ^* $(n \geqslant 1)$ is said to form a *bi-orthogonal* sequence if $f_n(\alpha_m) = 0$ $(m \neq n)$ and $f_n(\alpha_n) = 1$ $(n \geqslant 1)$. I shall also say that the sequence (f_n) is *ortho-normal* to the sequence (α_n). If S is any sub-set of Γ, I shall denote by $L(S)$ the closed linear extension of S: that is, the smallest closed linear sub-space of Γ containing S.

THEOREM 5. *Given a sequence (α_n) of Γ, a necessary and sufficient condition that there exists a sequence (f_n) of Γ^* ortho-normal to (α_n) is*

that, for each $i \geqslant 1$, $\alpha_i \notin L(\alpha_n, n \neq i)$. *If this condition is satisfied and* d_i *is the distance of* α_i *from* $L(\alpha_n, n \neq i)$, *then for each* $R > A(1/d_i)$ *we can find a functional* $f_i \in \Gamma^*$ *such that*

$$\text{(i)} \quad f_i(\alpha_n) = 0 \ (n \neq i) \quad \textit{and} \quad f_i(\alpha_i) = 1;$$

and

$$\text{(ii)} \quad |f_i(\alpha)| \leqslant |\alpha; R|/d_i \ \textit{for all} \ \alpha \in \Gamma.$$

Proof. That the condition in the theorem is sufficient and that (i) and (ii) are true follow immediately from Theorem 3. Conversely, if there is a functional $f_i \in \Gamma^*$ such that $f_i(\alpha_i) = 1$, $f_i(\alpha_n) = 0$, $n \neq i$, then obviously $\alpha_i \notin L(\alpha_n, n \neq i)$ since otherwise $f_i(\alpha_i)$ will also be zero.

5.1. Theorem 5 could be stated in an apparently more general form as follows.

THEOREM 6. *Given a sequence* (α_n) *of* Γ, *a necessary and sufficient condition that a sequence* (f_n) *of* Γ^* *exists ortho-normal to* (α_n) *is, that for each* $i \geqslant 1$, $\alpha_i \notin L(\alpha_n, n \geqslant i+1)$.

Proof. By Theorem 3, there is a sequence (g_n) of Γ^* such that $g_i(\alpha_i) = 1$, $g_i(\alpha_n) = 0$ $(n \geqslant i+1)$. Now we can construct the sequence (f_n) from (g_n) as follows. For each i, choose $(t_{i1}, t_{i2},..., t_{ii})$ from the equations

$$\left.\begin{aligned} t_{i1}g_1(\alpha_n)+t_{i2}g_2(\alpha_n)+...+t_{ii}g_i(\alpha_n) &= 0 \quad (n = 1,2,..., i-1) \\ t_{i1}g_1(\alpha_i)+t_{i2}g_2(\alpha_i)+...+t_{ii}g_i(\alpha_i) &= 1 \end{aligned}\right\}. \tag{10}$$

The determinant of the equations (10) is unity. So $(t_{i1},..., t_{ii})$ is uniquely determined by (10). The functional f_i can now be defined as

$$f_i = t_{i1}g_1+t_{i2}g_2+...+t_{ii}g_i.$$

It is easily verified by the choice of (g_i) and (10) that

$$f_i(\alpha_i) = 1, \qquad f_i(\alpha_n) = 0 \quad (n \neq i).$$

So the condition of Theorem 6 is sufficient. That it is necessary follows as in Theorem 5.

5.2. *Uniqueness of bi-orthogonal sequences.* Let (α_n) be a sequence of Γ satisfying the conditions of Theorem 5 or 6. If $L(\alpha_n) = \Gamma$, then the ortho-normal set (f_n) (whose existence is affirmed by these theorems) is unique. For, if there is another set (g_n), then, for each i, $f_i - g_i$ vanishes over the whole of $L(\alpha_n) = \Gamma$, and so $f_i = g_i$.

5.3. *Illustration.* Let

$$\alpha_n = z^n\beta_n(z), \qquad \beta_n(0) = 1 \quad (n \geqslant 0.)$$

Then for each i, $\alpha_i \notin L(\alpha_n, n \geqslant i+1)$, and the distance of α_i from

$L(\alpha_n, n \geqslant i+1)$ is not less than 1. So, by Theorem 3, there exists a $g_i \in \Gamma^*(R)$ such that

(i) $g_i(\alpha_i) = 1$, $g_i(\alpha_n) = 0$ for $n \geqslant i+1$, and

(ii) $|g_i(\alpha)| \leqslant |\alpha; R|$ for each $R > 1$.

Hence using (10) we see that, for each $R > 1$, there is a sequence (f_n) belonging to $\Gamma^*(R)$ ortho-normal to (α_n).

6. Bases in Γ

A sequence (α_n) of Γ is said to form a *base* in Γ if each $\alpha \in \Gamma$ could be *represented uniquely* in the form

$$\alpha = \sum_{n=1}^{\infty} t_n(\alpha)\alpha_n \tag{11}$$

where $\{t_n(\alpha)\}$ are complex numbers uniquely determined by α. In a Banach space (i.e. a complete normed linear space) it is known that each coefficient in (11) is a continuous linear functional [(B), 111]. I am going to prove that this result is true in the space Γ also.

THEOREM 7. *Let (α_n) be a sequence of elements of Γ forming a base. Let $f_n(\alpha) = t_n(\alpha)$ where $t_n(\alpha)$ is defined by (11). Then the sequence $\{f_n\}$ belongs to Γ^* and forms an ortho-normal set to (α_n).*

Proof. By the definition of a base, it follows that, for each n, $f_n(\alpha) = t_n(\alpha)$ is linear, $f_n(\alpha_m) = 0$ if $n \neq m$, and $f_m(\alpha_m) = 1$. So to prove the theorem it is enough to show that each $f_n \in \Gamma^*$. For this purpose, I define a new distance $|\alpha-\beta|_1$ in Γ where

$$|\alpha|_1 = \underset{1 \leqslant n < \infty}{\text{upper bound}} |t_1(\alpha)\alpha_1 + \ldots + t_n(\alpha)\alpha_n|. \tag{12}$$

Since the representation (11) is unique, we see that $|\alpha-\beta|_1$ is, in fact, a metric in Γ. I shall prove that Γ is complete under this new metric. Let (β_p) be a sequence in Γ such that $|\beta_p-\beta_q|_1 \to 0$ as $p, q \to \infty$. From (12) it follows that, for each i, $|[t_i(\beta_p)-t_i(\beta_q)]\alpha_i| \to 0$ as $p, q \to \infty$. Since no α_i is identically zero, we see that for each i, $t_i(\beta_p) \to t_i$ (say) as $p \to \infty$. Again, given ϵ, we can find p_0 so that

$$\left|\sum_{i=1}^{n} [t_i(\beta_p)-t_i(\beta_q)]\alpha_i\right| \leqslant \epsilon \quad \text{if} \quad p, q \geqslant p_0, \tag{13}$$

for *all* $n \geqslant 1$. Letting $q \to \infty$ in (13) we get

$$\left|\sum_{i=1}^{n} t_i(\beta_p)\alpha_i - \sum_{i=1}^{n} t_i\,\alpha_i\right| \leqslant \epsilon \quad \text{if} \quad p \geqslant p_0. \tag{14}$$

Taking $p = p_0$ in (14) and noting that (11) converges for $\alpha = \beta_{p_0}$, we see that we can find n_0 so that

$$\left|\sum_{i=1}^{n} t_i \alpha_i - \sum_{i=1}^{m} t_i \alpha_i\right| \leqslant 3\epsilon \quad \text{if} \quad n, m \geqslant n_0. \tag{15}$$

Since Γ is complete, (15) shows that $\sum_1^\infty t_i \alpha_i$ converges to an element $\beta \in \Gamma$. Since (α_n) is a base, it follows that $t_i = t_i(\beta)$. Hence from (14) we get that $|\beta_p - \beta|_1 \leqslant \epsilon$ if $p \geqslant p_0$. Hence Γ is complete under the new metric. From (11) and (12) it is easily seen that $|\alpha| \leqslant |\alpha|_1$ so that the topology under the metric $|\alpha-\beta|$ is stronger than that under $|\alpha-\beta|_1$. Since Γ is complete under both metrics, it follows from a known theorem [(B), 41, Theorem 6] that the two metrics are equivalent. So, if $|\beta_p| \to 0$ as $p \to \infty$, then $|\beta_p|_1 \to 0$ as $p \to \infty$. This latter implies, by (12), that $t_i(\beta_p) \to 0$ as $p \to \infty$ for each fixed i, since no α_i is identically zero. Hence $f_i(\alpha) = t_i(\alpha)$ belongs to Γ^*. This completes the proof† of Theorem 7.

6.1. *Examples of bases.* (i) Let $p_n(z)$ be a polynomial of degree n such that the coefficient of z^n is unity and the remaining coefficients form a bounded set as n varies $(n = 0, 1,...)$. Then it is known‡ that every integral function $\alpha = \alpha(z)$ could be expanded in the form

$$\alpha = \sum_0^\infty t_n p_n(z),$$

where the series converges uniformly in any finite circle. So $\alpha_n = p_n(z)$ $(n = 0, 1,...)$ will form a base provided that the above expansion is unique (this is not always the case). For instance, $p_0 = 1$, $p_1 = z-1$, $p_n = z^n - z^{n-1}$ $(n \geqslant 2)$ will form a base as may be verified by direct calculation.

(ii) Let $(\alpha_0, \alpha_1,..., \alpha_p)$ be $p+1$ integral functions such that the determinant formed by the coefficients of z^n $(n = 0, 1,..., p)$ does not vanish. If we set $\alpha_n = z^n$ for $n > p$, it is easily verified that (α_n) $(n \geqslant 0)$ form a base in Γ.

(iii) I next prove the following result.

THEOREM 8. *Let (α_n) $(n \geqslant 0)$ be a sequence of elements of Γ defined by*

$$\alpha_n = \alpha_n(z) = z^n + \sum_{p=n+1}^\infty a_{np} z^p. \tag{16}$$

† If (α_n) is a base, we have obviously $L(\alpha_n) = \Gamma$ and so, by § 5.2, the orthonormal set $\{t_i(\alpha)\}$ is unique.

‡ J. M. Whittaker, *Interpolatory Function Theory* (Cambridge Tracts, No. 33, 15, Theorem 2).

Let $\alpha_n(z)-z^n = \beta_n(z)$. If $|\beta_n| \to 0$ as $n \to \infty$, then the sequence (α_n) will form a base in Γ.

Proof. Let $\alpha = \sum\limits_0^\infty a_n z^n \in \Gamma$. Define (t_n) $(n \geqslant 0)$ by the equations

$$t_0 = a_0, \qquad t_n + a_{n-1,n} t_{n-1} + \ldots + a_{0n} t_0 = a_n \quad (n \geqslant 1). \tag{17}$$

The equations (17) determine (t_n) uniquely, and $t_n = 0$ for all n if and only if $a_n = 0$ for all n. Since $|\beta_n| \to 0$, we can find some $k > 1$ such that $|a_n| \leqslant k$, $|a_{np}| \leqslant k$ for all n and for all p under consideration. I shall prove by induction that

$$|t_n| \leqslant k(k+1)^n \quad (n \geqslant 0). \tag{18}$$

Obviously (18) is true for $n = 0$. Suppose it has been proved for $n = 0, 1, \ldots, p-1$. Then from (17) we get

$$|t_p| \leqslant k + k\{k + k(k+1) + \ldots + k(k+1)^{p-1}\}$$

$$= k + k^2 \frac{(k+1)^p - 1}{k} = k(k+1)^p.$$

So (18) is true. Using (18) I shall next prove that $|t_n|^{1/n} \to 0$ as $n \to \infty$. Let $\epsilon > 0$ be given. Since α is an integral function and $|\beta_n| \to 0$ as $n \to \infty$, by (2), we can find n_0 so that

$$\left.\begin{aligned} &|a_n| \leqslant \epsilon^n \\ &|a_{np}| \leqslant \epsilon^p \quad (p \geqslant n+1) \end{aligned}\right\} \quad \text{if } n \geqslant n_0. \tag{19}$$

For fixed n_0, $\beta_0, \beta_1, \ldots, \beta_{n_0-1}$ are a finite number of integral functions, and so we can choose $N \geqslant n_0$ such that

$$|a_{in}| \leqslant \epsilon^n \quad (i = 0, 1, \ldots, n_0-1;\; n \geqslant N). \tag{20}$$

Using (18), (19), and (20) in (17), we see that, if $n \geqslant N$,

$$|t_n| \leqslant \epsilon^n\{1 + k(k+1)^{n-1} + k(k+1)^{n-2} + \ldots + k\}$$

$$= \epsilon^n \left\{1 + k\frac{(k+1)^n - 1}{k}\right\}$$

$$= \epsilon^n (k+1)^n.$$

This inequality shows that $|t_n|^{1/n} \to 0$ as $n \to \infty$. Now, given any $R > 0$, it is easily verified, by using the fact that $|\beta_n| \to 0$ as $n \to \infty$, that

$$|\alpha_n(z)| = O(R^n) \quad \text{as } n \to \infty, \quad |z| \leqslant R.$$

Hence $\sum\limits_0^\infty t_n \alpha_n(z)$ converges uniformly in any finite circle to an integral

function $\beta(z) = \sum_0^\infty b_n z^n$. Using the usual formula for b_n, namely,

$$2\pi b_n = \int_{|z|=R} \frac{\beta(z)}{z^{n+1}}\,dz,$$

we easily verify that $b_n = a_n$ in virtue of (17). Hence $\beta(z) = \alpha(z)$. So, by (I), Theorem 3, we see that

$$\alpha = \sum_0^\infty t_n \alpha_n \quad \text{in } \Gamma,$$

and the representation is unique since $t_n = 0$ if and only if $a_n = 0$ for all n. Hence (α_n) form a base. For instance, if $\gamma(z)$ be any fixed integral function, we can take

$$\alpha_n = \alpha_n(z) = z^n + \left(\frac{z}{n}\right)^{n+1} \gamma\left(\frac{z}{n}\right).$$

(iv) The simplest base is (z^n), or more generally, $(pz+q)^n$ $(p \neq 0;$ $n \geqslant 0)$. It can easily be proved that *there exists no base of the form* (α^n) $(n \geqslant 0)$ *unless* $\alpha = pz+q$. To see this, suppose that α is either a polynomial of degree greater than one or a transcendental integral function. Suppose (α^n) forms a base. Then we must have

$$z \equiv \sum_0^\infty t_n [\alpha(z)]^n, \tag{21}$$

for proper (t_n) and for all z. Now we can find a complex number a and two *distinct* points z_1 and z_2 such that $\alpha(z_1) = \alpha(z_2) = a$. Using this in (21) we get

$$z_1 = z_2 = \sum_0^\infty t_n a^n$$

which is a contradiction. *Hence* (α^n) $(n \geqslant 0)$ *can form a base if and only if* $\alpha = pz+q$ $(p \neq 0)$.

(v) Let (α_n) be a base in Γ. Let U be an automorphism of Γ. Then $\{U(\alpha_n)\}$ will also form a base, as is evident from the definitions of a base and automorphism. If (t_n) is the unique ortho-normal set associated with (α_n), the ortho-normal set (f_n) associated with $\{U(\alpha_n)\}$ is defined by the relations $f_n(\alpha) = t_n[U^{-1}(\alpha)]$. For instance, let (p_n) be a sequence of complex numbers such that $\{|p_n|^{1/n}\}$ and $\{|p_n|^{-1/n}\}$ both form bounded sequences. Then

$$U(\alpha) = \sum_0^\infty p_n a_n z^n, \qquad \alpha = \sum_0^\infty a_n z^n$$

will define an automorphism of Γ since we can find $k > 0$ such that $k^{-1}|\alpha| \leqslant |U(\alpha)| \leqslant k|\alpha|$. So, if

$$(\alpha_n), \qquad \alpha_i = \sum_{n=0}^\infty a_{in} z^n$$

form a base, then so does the sequence $\{U(\alpha_n)\}$ where

$$U(\alpha_i) = \sum_{n=0}^{\infty} p_n a_{in} z^n.$$

7. Conclusion

It is easy to verify that Γ is a *convex* linear metric space. So results similar to Theorems 3 and 4 can be formulated in terms of pseudo-norms.† The interest of these theorems in this paper is that they have been formulated directly in terms of the metric and topology of Γ.

† See J. V. Wehausen, *Duke Math. J.* 4 (1938), 163.

ON A THEOREM OF BURNSIDE

By K. A. HIRSCH (*Newcastle-upon-Tyne*)

[Received 28 February 1949]

BURNSIDE proved* a remarkable relation between the (odd) order N of a group and the number r of its conjugate sets:

$$N \equiv r \pmod{16}.$$

This congruence is derived as an application of the theory of representations and of group characters.

In the present paper I prove by an elementary method a more general result which includes Burnside's as a special case.

THEOREM. *Let $N = p_1^{\alpha_1} p_2^{\alpha_2} \ldots p_k^{\alpha_k}$ be the order of a group G, the p's being primes, r the number of its conjugate sets, d the greatest common divisor of the numbers*

$$p_i^2 - 1 \quad (i = 1, 2, \ldots, k).$$

Then $\qquad N \equiv r \pmod{2d}$ *if N is odd,*

and $\qquad N \equiv r \pmod{3}$ *if N is even and $(N, 3) = 1$.*†

Proof. Denote the elements of the group by $X, Y, \ldots$, the unit element by 1.

Consider the equation

$$X^{-1}Y^{-1}XY = 1 \tag{1}$$

in G. If X lies in a given set of conjugates with, say, h elements, then the number of elements Y which are permutable with X is N/h. Hence the total number of solutions in G of the equation (1) is

$$\sum_{\rho=1}^{r} h_\rho N/h_\rho = Nr.$$

Among the solutions occurs $X = 1$, $Y = 1$. For all other solutions consider the group $\{X, Y\}$ generated by X and Y. X and Y have unique representations in G of the form

$$X = X_1 X_2 \ldots X_k, \qquad Y = Y_1 Y_2 \ldots Y_k,$$

where the X_i and Y_i have powers of p_i as orders. Since X and Y are permutable, we obtain the direct decomposition into Sylow-groups

$$\{X, Y\} = \{X_1, Y_1\} \times \{X_2, Y_2\} \times \ldots \times \{X_k, Y_k\}.$$

* Burnside, *Theory of Groups*, 2nd edition (Cambridge 1911), 295.

† This result can also be obtained by an extension of Burnside's original argument.

The number of ways in which a given group $\{X, Y\}$ arises from the solutions of equation (1) is, then, equal to the product of the number of ways in which its various Sylow-groups arise.

Now $\{X_i, Y_i\}$ can either be a cyclic group of order p_i^m, say, or an Abelian group of type (p_i^m, p_i^n) $(m \geqslant n)$. In the former case the number of ways in which the group can be generated by two of its elements is

$$p_i^{2m}-p_i^{2m-2} = p_i^{2m-2}(p_i^2-1) \equiv 0 \,(\mathrm{mod}\, d);$$

for either X_i or Y_i must be of order p_i^m, and of the possible p_i^{2m} pairs we have to rule out only those in which the orders of both X_i and Y_i are less than p_i^m.

In the latter case we distinguish between $m = n$ and $m > n$.

$[m = n]$: X_i and Y_i must be independent generators of the group. This yields

$$(p_i^{2m}-p_i^{2m-2})[(p_i^{2m}-p_i^{2m-2})-(p_i^m-p_i^{m-1})]$$

choices, and this number is obviously a multiple of $(p_i^2-1)(p_i-1)$ and divisible by d, and, for odd p_i, by $2d$.

$[m > n]$: **Either** X_i must be of order p_i^m, and Y_i of order p_i^n relative to the group $\{X_i\}$;

or X_i is of order less than p_i^m; then Y_i must be of order p_i^m, and X_i of order p_i^n relative to the group $\{Y_i\}$.

This yields $\phi(p_i^m)\, p_i^n\, \phi(p_i^n)\, (p_i^m+p_i^{m-1})$ choices and again this number is a multiple of

$$(p_i-1)^2(p_i+1) = (p_i^2-1)(p_i-1).$$

In all cases, then, the numbers of solutions of equation (1) which are different from $X = 1$, $Y = 1$ are multiples of (p_i^2-1), and hence of d. So we obtain $Nr \equiv 1 \,(\mathrm{mod}\, d)$. Now, since $p_i^2 \equiv 1 \,(\mathrm{mod}\, d)$, we have $N^2 \equiv 1 \,(\mathrm{mod}\, d)$, and therefore $N \equiv r \,(\mathrm{mod}\, d)$. In particular, if N is even but not divisible by 3, then $p_i^2-1 \equiv 0 \,(\mathrm{mod}\, 3)$ and $d = 3$.

This proves the second part of the theorem.

From now on I assume that N is odd.

Let $\qquad N^2 = 1+ld, \qquad Nr = 1+l'd.$

Then $\quad N^2r = r+lrd = N+Nl'd, \quad$ and $\quad N-r = (lr-l'N)d,$

and, since N and r are odd, we can infer $N \equiv r \,(\mathrm{mod}\, 2d)$ if and only if

$$l \equiv l' \,(\mathrm{mod}\, 2).$$

Let 2^β be the highest power of 2 which divides d, and put

$$p_i^2 = 1+l_i\, 2^\beta.$$

We have then

$$N^2 = \prod_{i=1}^{k} [1+l_i 2^\beta]^{\alpha_i} = \prod_{i=1}^{k} [1+\alpha_i l_i 2^\beta + \ldots] = 1+\sum_{i=1}^{k} \alpha_i l_i 2^\beta + \ldots$$

and

$$(N^2-1)/2^\beta \equiv \sum_{i=1}^{k} \alpha_i l_i \pmod 2.$$

We see, therefore, that l is congruent (mod 2) to the number of primes p_i with the properties (i) l_i odd, (ii) α_i odd.

I consider now the distribution of the solutions of equation (1) in multiples of $2^{\beta+1}$. The solutions arising from Abelian groups of type (p_i^m, p_i^n), as we have seen, always occur in multiples of $2^{\beta+1}$. The same is true of the cyclic groups $\{X, Y\}$ whose order contains more than one prime factor; and of the cyclic groups of prime power order for those primes p_i whose l_i is even. It remains to investigate the cyclic groups whose order is a power of p_i with odd l_i.

We now require the simple

Lemma. *The total number of cyclic sub-groups $\neq 1$ in a group of order p^n is congruent to n* (mod 2).

For every element other than 1 generates some cyclic p-group, and, if there are λ_j groups of order p^j, they give rise to $\lambda_j \phi(p^j)$ generating elements. Hence

$$p^n - 1 = \sum_j \lambda_j \phi(p^j) = \sum_j \lambda_j p^{j-1}(p-1),$$

$$p^{n-1}+p^{n-2}+\ldots+p+1 = \sum_j \lambda_j p^{j-1},$$

and we obtain $n \equiv \sum_j \lambda_j \pmod 2$, where the right-hand side gives the total number of cyclic sub-groups not equal to 1.

Now, since N is odd, the number of different Sylow-groups belonging to $p_i^{\alpha_i}$ is odd; and, among these, the number into which a given cyclic group of order p_i^n enters is odd. Hence we can restrict ourselves (mod $2^{\beta+1}$) to the total number of solutions which arise from cyclic sub-groups of *one* Sylow-group.

The number of cyclic sub-groups of one Sylow-group of order $p_i^{\alpha_i}$ is, however, by the preceding lemma congruent to $\alpha_i \pmod 2$. Hence l' is congruent (mod 2) to the number of primes p_i with the properties (i) l_i odd, (ii) α_i odd.

Thus $l \equiv l' \pmod 2$. This concludes the proof of the theorem. The modulus $2d$ is the best possible, as simple examples show. It may be noted that, if N is odd and not divisible by 3, the modulus of the congruence is always a multiple of 48. It may be considerably higher; e.g. if N is divisible by 29, 41, and 71 only, then $N \equiv r \pmod{1680}$.

THE CHARACTERIZATION OF GENERALIZED CONVEX FUNCTIONS

By F. F. BONSALL (*Newcastle*)

[Received 1 March 1949]

1. Introduction

A CONVEX curve may be defined as a curve which lies on or below the line segment joining any two points on it. We may describe this situation by saying that a convex function is 'majorized' by the solutions of the differential equation $d^2y/dx^2 = 0$. In the present paper, convex functions are generalized by considering functions majorized by the solutions of a more general differential equation.

2. Let $L(y) = 0$ denote the second-order linear differential equation

$$L(y) \equiv d^2y/dx^2+p_1(x)\,dy/dx+p_2(x)y = 0. \tag{2.1}$$

I consider, throughout, an interval (a,b) such that $L(y) = 0$ has a *unique* solution, continuous in (a,b), taking any given real values y_1 and y_2 at any two given x_1 and x_2 within (a,b). I suppose $p_1(x)$ and $p_2(x)$ continuous in (a,b), and differentiable whenever this is required.

3. Definition of sub-(L) functions

A real function $f(x)$, defined in $a < x < b$, is said to be 'sub-(L) in (a,b)' if

$$f(x) \leqslant F(f,x_1,x_2;x) \tag{3.1}$$

for every x, x_1, x_2 such that

$$a < x_1 < x < x_2 < b,$$

$F(f,x_1,x_2;x)$ being the solution of (2.1) taking the values $f(x_1)$ and $f(x_2)$ at x_1 and x_2.

Further, $f(x)$ is said to be *strictly* sub-(L) if the strict inequality holds in (3.1). 'Super-(L)' functions are defined in a similar way by reversing the inequality (3.1).

4. The principal results are the following theorems on the characterization of sub-(L) functions.

THEOREM 4. *If $f(x)$ is sub-(L) in (a,b), then f has a second derivative p.p. in (a,b), and $L(f) \geqslant 0$ at each point where the second derivative exists.*

THEOREM 6. *If $f(x)$ has a continuous second derivative and $L(f) \geqslant 0$ in (a,b), then f is sub-(L) in (a,b). If the strict inequality holds, f is strictly sub-(L).*

5. By means of these characterization theorems, it is possible to use sub-(L) functions as an analytical tool in a manner somewhat similar to the use of convex functions. The applications of the theory in the present paper are not systematic, but are intended merely to indicate some of the possibilities. Some determinantal inequalities are obtained, the behaviour of the mean-value function $M_l(a,q)$, defined in (19.1), is briefly investigated, and a method is outlined for estimating solutions of differential equations which cannot be explicitly solved.

It is evident that sub-harmonic functions can be similarly generalized, a general elliptic partial differential equation of the second order replacing the Laplace equation as 'majorizing' equation.

Sub-(L) functions are a sub-class of the functions 'convex in ϕ and ψ', considered by Valiron [**1**]. They are, *a fortiori*, a sub-class of the sub-(F) functions considered by Beckenbach [**2**] and Beckenbach and Bing [**3**]. On the other hand, ordinary convex functions and the sub-trigonometrical functions of Pólya [**4**] are particular cases of sub-(L) functions.

[*Added* 18 *Sept.* 1949.] A part of my results has been anticipated by Peixoto in a recent paper [**8**]. His theory is in one sense more general, in that it applies to a class of non-linear differential inequalities. On the other hand, his proofs require the functions to have continuous second derivatives.

6. Notation

The following notation is adopted:

(i) $F(f,x_i,x_j;x)$ is that solution of $L(y) = 0$ for which

$$F(f,x_i,x_j;x_i) = f(x_i), \qquad F(f,x_i,x_j;x_j) = f(x_j). \tag{6.1}$$

Where no ambiguity can arise, I write $F_{ij}(x)$ for $F(f,x_i,x_j;x)$.

(ii) $F_{0,h}(x)$ is the solution $F(f,x_0,x_0+h;x)$.

(iii) (α,β) and $\langle\alpha,\beta\rangle$ denote open and closed intervals respectively.

(iv) $\bar{L}(y) = 0$ is the adjoint equation of $L(y) = 0$,

i.e. $$\bar{L}(y) \equiv d^2y/dx^2 - p_1(x)\,dy/dx + [p_2(x) - p_1'(x)]y = 0. \tag{6.2}$$

7. Some preliminary lemmas

The fundamental role in the proof of the properties of sub-(L) functions is played by the uniqueness condition in § 2. Whenever I write 'by uniqueness', it is to this condition that I refer.

LEMMA 1. *If $f(x)$ is sub-(L) in (a,b), then*

$$f(x) \geqslant F_{12}(x) \tag{7.1}$$

for x in (a,x_1) or (x_2,b).

COROLLARY. *A function sub-(L) in (a,b) is bounded above and below in every closed sub-interval of (a,b).*

Suppose on the contrary that, say,

$$F_{12}(x_3) > f(x_3) = F_{13}(x_3) \quad (x_1 < x_2 < x_3). \tag{7.2}$$

Then, since
$$F_{12}(x_2) = f(x_2) \leqslant F_{13}(x_2), \tag{7.3}$$

we infer, by continuity of $F_{12}(x)-F_{13}(x)$, that there exists an x' such that $x_2 \leqslant x' < x_3$ and $F_{12}(x') = F_{13}(x')$. Hence, by uniqueness,

$$F_{12}(x) \equiv F_{13}(x),$$

which contradicts (7.2).

Finally, $f(x)$ is bounded above in $\langle x_1, x_2 \rangle$, by (3.1), and it is bounded below since
$$f(x) \geqslant F_{45}(x) \quad \text{in} \quad \langle x_1, x_2 \rangle, \tag{7.4}$$

where $a < x_4 < x_5 < x_1$.

8. The following lemma on the adjoint equation is presumably known.

LEMMA 2. *If L satisfies the uniqueness condition in (a,b), so does $\bar{L}$.*

For, we have identically

$$\int\limits_{x_1}^{x_2} [v\,L(u)-u\bar{L}(v)]\,dx = [vu'-uv'+p_1\,uv]_{x_1}^{x_2}. \tag{8.1}$$

Suppose the lemma false, i.e. we may suppose that there exists a solution $v(x)$ of $\bar{L}(y) = 0$, not identically zero but satisfying $v(x_1) = v(x_2) = 0$. Then choosing $u(x)$ to be that solution of $L(y) = 0$ for which $u(x_1) = 1$, $u(x_2) = 0$, we obtain $v'(x_1) = 0$, and thus $v(x) \equiv 0$.

9. LEMMA 3. *Given $a < x_1 < x_0 < x_2 < b$, let $u(x)$, $v(x)$ be the solutions of $\bar{L}(y) = 0$ satisfying*

$$u(x_1) = 0, \qquad u(x_0) = -1; \qquad v(x_0) = -1, \qquad v(x_2) = 0. \tag{9.1}$$

Then
$$u'(x_1) < 0 \quad v'(x_2) > 0, \quad v'(x_0)-u'(x_0) > 0. \tag{9.2}$$

This lemma is proved by a simple application of uniqueness. It is, of course, equally true if u, v are the solutions of $L(y) = 0$ satisfying (9.1), but the present notation is required in the proof of Theorem 5.

10. Continuity and derivability

THEOREM 1. *If f is sub-(L) in (a,b), it is continuous and has a right-hand derivative f'_+ and a left-hand derivative f'_- in (a,b). Also,*

$$f'_+ \geqslant f'_- \tag{10.1}$$

and indeed $f'_+ = f'_-$ except possibly in an enumerable set of points.

COROLLARY. *If for any x_0 in (a,b), $y = v(x)$ is a solution of $L(y) = 0$ such that*

(i) $v(x_0) = f(x_0)$,

(ii) $f'_-(x_0) \leqslant v'(x_0) \leqslant f'_+(x_0)$,

then $y = v(x)$ is a 'curve of support' for the curve $y = f(x)$;

i.e.
$$v(x) \leqslant f(x) \tag{10.2}$$
for all x in (a,b).

Proof. Let x_0, x_0+h be any points in (a,b). Then
$$F_{0,h}(x) = A(h)e_1(x)+B(h)e_2(x) \tag{10.3}$$
where $e_1(x)$, $e_2(x)$ are two linearly independent solutions of (2.1) in (a,b).

Here
$$A(h) = \frac{f(x_0)e_2(x_0+h)-f(x_0+h)e_2(x_0)}{e_1(x_0)e_2(x_0+h)-e_1(x_0+h)e_2(x_0)}, \tag{10.4}$$
and there is a similar expression for $B(h)$.

Now, if $0 < h < h'$,
$$F_{0,h}(x_0+h) = f(x_0+h) \leqslant F_{0,h'}(x_0+h), \tag{10.5}$$
and hence, by uniqueness,
$$F_{0,h}(x) \leqslant F_{0,h'}(x) \quad \text{for} \quad x_0 \leqslant x < b. \tag{10.6}$$
Thus $F_{0,h}(x)$ decreases as h decreases for every x in (x_0,b).
Also $F'_{0,h'}(x_0) \geqslant F'_{0,h}(x_0)$ and so, again by uniqueness,
$$F_{0,h'}(x) \leqslant F_{0,h}(x) \quad \text{in } (a,x_0). \tag{10.7}$$
Thus $F_{0,h}(x)$ increases as h decreases for every x in (a,x_0). It is readily deduced from Lemma 1 that the family of functions $F_{0,h}(x)$ with $h > 0$ is uniformly bounded below in (x_0,b) and above in (a,x_0).

We have now proved that
$$F_{0,+}(x) = \lim_{h\to+0} F_{0,h}(x) \tag{10.8}$$
exists in (a,b). Also
$$F_{0,+}(x) \leqslant f(x) \quad \text{in } (a,b). \tag{10.9}$$
This follows at once from Lemma 1 in (a,x_0), and it follows in (x_0,b) by letting $h \to +0$ in
$$F_{0,+}(x) \leqslant F_{0,h}(x) \leqslant f(x) \quad (x > x_0+h). \tag{10.10}$$
Similarly, $F_{0,-}(x)$ $[= \lim_{h\to-0} F_{0,h}(x)]$ exists and satisfies
$$F_{0,-}(x) \leqslant f(x) \quad \text{in } (a,b). \tag{10.11}$$

Since, for $h < 0 < h'$, $F_{0,h}(x) \leqslant F_{0,h'}(x)$ in (x_0, b), we see that

$$F_{0,-} \leqslant F_{0,+} \quad \text{in } (x_0, b), \quad \text{and} \quad F_{0,-} \geqslant F_{0,+} \quad \text{in } (a, x_0). \tag{10.12}$$

Denoting $e_i(x_0)$, $e_i'(x_0)$ by e_i, e_i', we may write

$$A(h) = \frac{f(x_0)e_2' - e_2\Delta_h + O(h)}{e_1e_2' - e_2e_1' + O(h)}, \tag{10.13}$$

where $\Delta_h = [f(x_0+h) - f(x_0)]/h$, and $e_1e_2' - e_2e_1' \neq 0$. Thus, with a similar expression for $B(h)$,

$$F_{0,h}(x) = \frac{\Delta_h[e_1e_2(x) - e_2e_1(x)] + f(x_0)[e_2'e_1(x) - e_1'e_2(x)] + O(h)}{e_1e_2' - e_2e_1' + O(h)}. \tag{10.14}$$

Now $F_{0,h}(x)$ tends to a finite limit as $h \to +0$, and $e_1e_2(x) - e_2e_1(x)$ vanishes only at x_0. Hence, Δ_h tends to a finite limit f_+' as $h \to +0$. Similarly, f_-' exists. The continuity of $f(x)$ is now obvious. Further,

$$F_{0,+}(x) = \frac{f_+'[e_1e_2(x) - e_2e_1(x)] + f(x_0)[e_2'e_1(x) - e_1'e_2(x)]}{e_1e_2' - e_2e_1'}, \tag{10.15}$$

showing that $F_{0,+}(x)$ is the solution of $L(y) = 0$ satisfying

$$F_{0,+}(x_0) = f(x_0), \qquad F_{0,+}'(x_0) = f_+'(x_0). \tag{10.16}$$

Similarly, $F_{0,-}'(x_0) = f_-'(x_0)$, and so, by (10.12), $f_+' \geqslant f_-'$. Finally, it is well known that for any $f(x)$, the set of points where f_+' and f_-' exist and are different is enumerable [**6**].

The corollary follows at once from (10.9), (10.11), (10.16).

11. THEOREM 2. *If $f(x)$ is sub-(L) in (a, b), then f_+' and f_-' are bounded in any closed sub-interval $\langle x_1, x_2\rangle$ of (a, b).*

Proof. Given x_1, x_2 in (a, b), choose any x_0, x_3, x_4 such that

$$a < x_3 < x_1 \leqslant x_0 \leqslant x_2 < x_4 < b. \tag{11.1}$$

With our usual notation, we know that, for $x_1 \leqslant x < b$,

$$F_{13}(x) \leqslant F_{10}(x) \leqslant F_{12}(x) \tag{11.2}$$

(in the special case $x_0 = x_1$ we allow F_{10} to stand for $F_{1,+}$).

Denoting by $e_1(x)$, $e_2(x)$ the basic solutions of $L(y) = 0$, such that $e_1(x_1) = 1$, $e_1'(x_1) = 0$; $e_2(x_1) = 0$, $e_2'(x_1) = 1$, we find

$$F_{10}(x) = f(x_1)e_1(x) + \theta e_2(x), \tag{11.3}$$

where $\theta = F_{10}'(x_1)$.

By (11.2), we know that

$$F_{13}'(x_1) \leqslant \theta \leqslant F_{12}'(x_1). \tag{11.4}$$

Now $F'_{10}(x)$ $[=f(x_1)e'_1(x)+\theta e'_2(x)]$ is a continuous function of (x,θ), and hence, for (x,θ) in the closed rectangle

$$x_1 \leqslant x \leqslant x_2, \quad F'_{13}(x_1) \leqslant \theta \leqslant F'_{12}(x_1),$$

it attains finite upper and lower bounds, M_1 and m_1 say. In particular,

$$m_1 \leqslant F'_{10}(x_0) \leqslant M_1 \tag{11.5}$$

for all x_0 in $\langle x_1, x_2 \rangle$, and hence

$$f'_+(x_0) \geqslant f'_-(x_0) \geqslant F'_{10}(x_0) \geqslant m_1 \tag{11.6}$$

for x_0 in $\langle x_1, x_2 \rangle$.

Similarly, by considering the solutions F_{02}, F_{12}, F_{24} we have

$$f'_-(x_0) \leqslant f'_+(x_0) \leqslant M_2 \quad \text{in} \quad \langle x_1, x_2 \rangle. \tag{11.7}$$

12. THEOREM 3. *If f is sub-(L) in (a,b), then, given x_1, x_2 in (a,b), f is the difference of two functions continuous and convex in (x_1,x_2).*

COROLLARY. (i) *In any such (x_1,x_2), f is the integral of a function of bounded variation.*

(ii) *f has a second derivative p.p. in (a,b).*

Proof. Let $x_1 < x_0 < x_2$, and set $f_n(x) = f(x)+nx^2$, and

$$F_n(x) = F_{0,+}(x)+nx^2.$$

Then, by (10.9), $\quad F_n(x) \leqslant f_n(x) \quad (a < x < b) \tag{12.1}$

and $\quad F_n(x_0) = f_n(x_0), \qquad F'_n(x_0) = f'_{n,+}(x_0).$

Now

$$\begin{aligned} F''_n(x) &= F''_{0,+}(x)+2n \\ &= -[p_1(x)F'_{0,+}(x)+p_2(x)F_{0,+}(x)]+2n, \end{aligned}$$

where p_1, p_2 are the coefficients in L. In particular,

$$F''_n(x_0) = -[p_1(x_0)f'_+(x_0)+p_2(x_0)f(x_0)]+2n. \tag{12.2}$$

But, f'_+ being bounded and f, p_1, p_2 continuous in $\langle x_1, x_2 \rangle$, we can choose a sufficiently large n so that, for all x_0 in $\langle x_1, x_2 \rangle$, $F''_n(x_0) > 1$, say.

Hence, for $x_0 < x < x_0+\delta_0$ and sufficiently small positive δ_0,

$$F_n(x) > F_n(x_0)+(x-x_0)F'_n(x_0). \tag{12.3}$$

From (12.3) with (12.1) we have

$$f_n(x) > f_n(x_0)+(x-x_0)f'_{n,+}(x_0), \tag{12.4}$$

for $x_1 \leqslant x_0 \leqslant x_2$, $x_0 < x < x_0+\delta_0$.

To prove $f_n(x)$ convex in (x_1,x_2) it is sufficient to prove that (12.4) holds for all x_0 in (x_1,x_2) and x in (x_0,x_2). Suppose then that, for some x_0, (12.4) does not hold for all x in (x_0,x_2); then there is an $x' > x_0$ such that

$$f_n(x') = f_n(x_0)+(x'-x_0)f'_{n,+}(x_0). \tag{12.5}$$

Then $f_n(x)-f_n(x_0)-(x-x_0)f'_{n,+}(x_0)$ attains a maximum in (x_0, x'), but this is impossible since this function satisfies (10.1) and (12.4).

To complete the theorem it need only be remarked that nx^2 is a convex function. Corollary (i) is now an immediate deduction from the well-known theorem [7] that a continuous convex function is the integral of an increasing function, since the difference of two increasing functions is a function of bounded variation. Finally Corollary (ii) follows from (i).

13. Characterization of sub-(L) functions

The main theorems stated in § 4 may now be proved.

Proof of Theorem 4. The first part of the theorem has been proved already. Suppose that $f''(x_0)$ exists for some x_0 in (a,b). Since $f'(x_0)$ exists, we see, by (10.16), that

$$F_0(x) = \lim_{h\to 0} F_{0,h}(x)$$

is the solution of $L(y) = 0$ satisfying

$$F_0(x_0) = f(x_0), \qquad F'_0(x_0) = f'(x_0). \tag{13.1}$$

Thus, at $x = x_0$,

$$L(f) = L(f-F_0) = f''(x_0)-F''_0(x_0). \tag{13.2}$$

To complete the theorem, I now prove that $f''(x_0) \geqslant F''_0(x_0)$. In fact, writing $g(x) = F_0(x)-f(x)$, suppose that $g''(x_0) = \epsilon > 0$. Then

$$g'(x_0+h) = g'(x_0+h)-g'(x_0) > \tfrac{1}{2}\epsilon h$$

for all sufficiently small $h > 0$. Since $g''(x_0)$ exists, g has a first derivative in a neighbourhood of x_0, and so

$$\begin{aligned} g(x_0+h) = g(x_0+h)-g(x_0) &= hg'(x_0+\theta h) \quad (\theta > 0) \\ &> \tfrac{1}{2}\epsilon\theta h^2 > 0; \end{aligned} \tag{13.3}$$

but this contradicts the corollary to Theorem 1.

14. THEOREM 5. *Given $a < x_1 < x_0 < x_2 < b$, positive numbers α, β, γ, depending on x_0, x_1, x_2, and L, but not on $f(x)$, can be chosen so that*

(i) $\alpha f(x_0) = \beta f(x_1)+\gamma f(x_2)$ *if f is a solution of $L(y) = 0$,*
(ii) $\alpha f(x_0) \leqslant \beta f(x_1)+\gamma f(x_2)$ *if f is sub-(L).*

Proof. With the notation of Lemma 3, let

$$\alpha = v'(x_0)-u'(x_0), \qquad \beta = -u'(x_1), \qquad \gamma = v'(x_2). \tag{14.1}$$

Then α, β, γ are positive and independent of f. Suppose first that f has a continuous second derivative in (a,b). Then

$$\begin{aligned}\int_{x_1}^{x_0} uL(f)\,dx &= \int_{x_1}^{x_0} \{uL(f)-f\bar{L}(u)\}\,dx\\ &= [uf'-fu'+p_1uf]_{x_1}^{x_0}\\ &= -f'(x_0)-f(x_0)u'(x_0)-p_1(x_0)f(x_0)+f(x_1)u'(x_1).\end{aligned}$$

Combining this with a similar expression for $\int_{x_0}^{x_2} vL(f)\,dx$,

$$\int_{x_1}^{x_0} uL(f)\,dx+\int_{x_0}^{x_2} vL(f)\,dx = \alpha f(x_0)-\beta f(x_1)-\gamma f(x_2). \tag{14.2}$$

Theorem 5 (i) now follows at once from (14.2). Finally, suppose that f is any sub-(L) function, then

$$\begin{aligned}\alpha f(x_0) \leqslant \alpha F_{12}(x_0) &= \beta F_{12}(x_1)+\gamma F_{12}(x_2)\\ &= \beta f(x_1)+\gamma f(x_2).\end{aligned} \tag{14.3}$$

15. Proof of Theorem 6

Suppose, for example, that $L(f) > 0$ in (a,b). Then, given any x_0, x_1, x_2 for which $a < x_1 < x_0 < x_2 < b$, we see, since $u < 0$, $v < 0$, that

$$\begin{aligned}\alpha f(x_0) &< \beta f(x_1)+\gamma f(x_2)\\ &= \beta F_{12}(x_1)+\gamma F_{12}(x_2) = \alpha F_{12}(x_0).\end{aligned} \tag{15.1}$$

Since α is positive, $\qquad f(x_0) < F_{12}(x_0)$.

16. Non-homogeneous differential equations

The characterization of sub-(L) functions may be extended at once to non-homogeneous L. Let

$$L_1(y) \equiv d^2y/dx^2+p_1\,dy/dx+p_2\,y-q(x) \equiv L(y)-q(x), \tag{16.1}$$

and let $Q(x)$ be a particular solution of $L_1(y) = 0$. Then the general solution is of the form

$$y = Q(x)+A_{12}\,e_1(x)+B_{12}\,e_2(x), \tag{16.2}$$

where $e_1(x)$, $e_2(x)$ are basic solutions of $L(y) = 0$. Now, if f is sub-(L_1), we have, for x in (x_1, x_2),

$$f(x) \leqslant Q(x)+A_{12}\,e_1(x)+B_{12}\,e_2(x), \tag{16.3}$$

where $\qquad f(x_i) = Q(x_i)+A_{12}\,e_1(x_i)+B_{12}\,e_2(x_i) \quad (i = 1, 2)$,

i.e. $f(x)-Q(x)$ is sub-(L), and so, at any point where f has a second derivative, $L(f-Q) \geqslant 0$. Hence

$$L(f) \geqslant L(Q) = q(x), \quad \text{i.e.} \quad L_1(f) \geqslant 0.$$

The converse theorem is similarly established.

17. An integral inequality for sub-(L) functions

THEOREM 7. *If f is sub-(L), then, for $0 < h < h_0$*

$$f(x_0) \leqslant \frac{1}{2h}\left\{\int_{x_0-h}^{x_0} f\bar{L}(u)\,dx + \int_{x_0}^{x_0+h} f\bar{L}(v)\,dx\right\}, \tag{17.1}$$

where $u = \frac{1}{2}(x-x_0+h)^2$, $v = \frac{1}{2}(x-x_0-h)^2$, and h_0 depends on L, x_0 but not on f.

This theorem generalizes the known inequality for convex f:

$$f(x_0) \leqslant \frac{1}{2h}\int_{x_0-h}^{x_0+h} f(x)\,dx.$$

Proof. Let $F(x) = F_{0,+}(x)$ denote the solution of $L(y) = 0$ satisfying $F(x_0) = f(x_0)$, $F'(x_0) = f'_+(x_0)$. Then, by the corollary to Theorem 1,

$$f(x) \geqslant F(x) \quad \text{in} \quad (a,b).$$

Also,
$$\bar{L}(v) = 1-p_1(x-x_0-h)+\tfrac{1}{2}(p_2-p_1')(x-x_0-h)^2$$
$$> 0 \quad \text{in } (x_0, x_0+h) \text{ for sufficiently small } h.$$

Similarly, $\bar{L}(u) > 0$ in (x_0-h, x_0) for sufficiently small h. Hence,

$$\begin{aligned}\int_{x_0}^{x_0+h} f\bar{L}(v)\,dx &\geqslant \int_{x_0}^{x_0+h} F\bar{L}(v)\,dx\\ &= [v'F-vF'-p_1vF]_{x_0}^{x_0+h}\\ &= F(x_0)[-v'(x_0)+p_1(x_0)v(x_0)]+v(x_0)F'(x_0)\\ &= [h+\tfrac{1}{2}h^2p_1(x_0)]f(x_0)+\tfrac{1}{2}h^2F'(x_0).\end{aligned} \tag{17.2}$$

Similarly,

$$\int_{x_0-h}^{x_0} f\bar{L}(u)\,dx \geqslant [h-\tfrac{1}{2}h^2p_1(x_0)]f(x_0)-\tfrac{1}{2}h^2F'(x_0),$$

and the theorem now follows by addition.

APPLICATIONS

18. Some determinantal inequalities

It is known [5] that, if $x_1 < x_2 < x_3$ and $y_1 < y_2 < y_3$, then the determinant

$$|\exp(x_i y_j)| > 0 \quad (i, j = 1, 2, 3). \tag{18.1}$$

This result may be obtained at once by considering the condition that $e^{\gamma x}$ be sub-(L) when $L = (D-\alpha)(D-\beta)$. By considering in a similar manner the equation $(D-\alpha)^2y = 0$, it may be proved that, for

$$x_1 < x_2 < x_3 \quad \text{and} \quad \alpha \neq \beta,$$

$$\begin{vmatrix} e^{\beta x_1} & e^{\alpha x_1} & x_1 e^{\alpha x_1} \\ e^{\beta x_2} & e^{\alpha x_2} & x_2 e^{\alpha x_2} \\ e^{\beta x_3} & e^{\alpha x_3} & x_3 e^{\alpha x_3} \end{vmatrix} > 0. \tag{18.2}$$

Similar inequalities may be obtained for hypergeometric functions by considering their differential equations.

19. Improvement of known inequalities

I consider, for example, the mean-value function $M_t(a,q)$ defined by

$$M_t(a,q) = \Big\{\sum_{i=1}^{n} q_i\, a_i^t\Big\}^{1/t} = \Big\{\sum_{i=1}^{n} q_i\, e^{\alpha_i t}\Big\}^{1/t} \quad (q_i, a_i > 0;\ \textstyle\sum q_i = 1). \tag{19.1}$$

Writing $f(t) = \log M_t$, we obtain

$$f''(t)+\frac{2}{t}f'(t) = \frac{\sum q_j q_k(\alpha_j-\alpha_k)^2 e^{(\alpha_j+\alpha_k)t}}{t(\sum q_i e^{\alpha_i t})^2}$$

$$= \frac{1}{t}p(t), \quad \text{say.} \tag{19.2}$$

Now $p(t) > 0$ unless $\alpha_j = \alpha_k$ for all j, k. Hence we have

THEOREM 8. *If $L(y) = d^2y/dt^2+(2/t)\,dy/dt$, then $\log M_t(a,q)$ is sub-(L) for $t > 0$ and super-(L) for $t < 0$. The theorem is true in the strict sense unless all the a_i are equal.*

The general solution of $L(y) = 0$ is $y = A/t+B$, and, expressing the fact that $\log M_t$ is sub-(L) for $t > 0$, we obtain the known result [6] that $\log(M_t)^t$ is a convex function of t. However, stronger results may be obtained by considering more closely the behaviour of $p(t)$.

THEOREM 9. *With the notation of (19.2), suppose that $\lambda \leqslant p(t) \leqslant \mu$ for $0 < t_1 \leqslant t \leqslant t_2$, then*

$$-\tfrac{1}{2}\mu(t_2-t)(t-t_1) \leqslant \log(M_t)^t - \frac{t_2-t}{t_2-t_1}\log(M_{t_1})^{t_1} - \frac{t-t_1}{t_2-t_1}\log(M_{t_2})^{t_2}$$

$$\leqslant -\tfrac{1}{2}\lambda(t_2-t)(t-t_1). \tag{19.3}$$

Proof. If $f(t) = \log M_t$, we have at once

$$f''(t)+(2/t)f'(t)-\lambda/t \geqslant 0 \quad \text{in } (t_1,t_2), \tag{19.4}$$

so that f is sub-(L) in (t_1,t_2) when $L(y) = 0$ is

$$d^2y/dt^2+(2/t)\,dy/dt-\lambda/t = 0, \tag{19.5}$$

the general solution of which is $y = A/t+B+\frac{1}{2}\lambda t$. The theorem now follows by straightforward manipulation.

Moreover, it is evident that stronger theorems may be obtained by estimating $p(t)$ more closely. No doubt, all such results can be obtained by other methods, but sub-(L) functions provide a simple general technique for obtaining them.

20. Estimation of solutions of differential equations

The theory of sub-(L) functions may be used to obtain upper and lower estimates for the solutions of differential equations of type $y'' = F(x,y,y')$ which cannot be explicitly solved. To illustrate some possible techniques, I consider three examples.

(A) Let $f(x)$ be a solution of the non-linear equation

$$d^2y/dx^2+a\,dy/dx+by = q(x)\,(dy/dx)^2 \quad (a,\ b \text{ constant}), \tag{20.1}$$

which may be written $L(y) = q\,(dy/dx)^2$. Suppose that it is known that

(i) $m \leqslant [f'(x)]^2 \leqslant M$ for x in (x_1, x_2),

(ii) $q(x) \geqslant 0$ in (x_1, x_2).

Then $L_1(f) \geqslant 0$ and $L_2(f) \leqslant 0$, where

$$L_1(y) = L(y)-mq, \qquad L_2(y) = L(y)-Mq.$$

It follows that, if $e^{\alpha x}$ and $e^{\beta x}$ are basic solutions of $L(y) = 0$ and $Q(x)$ is a particular solution of $L(y) = q(x)$, then $f(x)$ is majorized by the functions $Ae^{\alpha x}+Be^{\beta x}+mQ(x)$, and minorized by the functions

$$Ae^{\alpha x}+Be^{\beta x}+MQ(x).$$

(B) Let $f(x)$ be a solution of the non-linear equation

$$d^2y/dx^2 = y^2+6x \tag{20.2}$$

which is defined in $(0, 1)$ and satisfies $f(0) = f(1) = 0$. Now $f^2 \geqslant 0$, so that f is sub-(L) for the equation $d^2y/dx^2 = 6x$, and therefore

$$f(x) \leqslant x^3-x < 0 \quad \text{in} \quad (0, 1).$$

We may obtain a sequence of majorizing functions $s_n(x)$ by iteration. Thus $f^2 \geqslant x^6-2x^4+x^2$, so that f is sub-(L) for the equation

$$d^2y/dx^2 = x^6-2x^4+x^2+6x,$$

and so on. More generally, we may prove that

$$f(x) \leqslant u_1(x)+u_2(x)+...+u_n(x) = s_n(x),$$

where $u_r(x)$ is a polynomial non-positive and convex in $(0, 1)$ *and vanishing at* 0, 1.

In fact, $s_n(x)$ is the solution of $d^2y/dx^2 = s_{n-1}^2+6x$, such that

$$s_n(0) = s_n(1) = 0.$$

Suppose that $u_r \leqslant 0$ in $(0,1)$ $(r = 1,..., n)$; then $u_{n+1} \leqslant 0$ in $(0,1)$. For, $s_n^2 = (u_n+s_{n-1})^2 = t_n+s_{n-1}^2$, where $t_n \geqslant 0$ in $(0,1)$. Now u_{n+1} is the solution of $d^2y/dx^2 = t_n$, vanishing at 0, 1, and it follows at once that u_{n+1} is non-positive and convex in $(0,1)$.

(C) *Sub-(L) properties of characteristic functions.*

Let $$L(y) = d(ky')/dx-ly, \qquad k(x), l(x) > 0; \tag{20.3}$$

and let $\{\phi_n\}$ be a complete system of characteristic functions corresponding to the positive increasing characteristic numbers $\{\lambda_n\}$, for the equation $[L+\lambda g]y = 0$, $g(x) > 0$, subject to Sturmian boundary conditions. Then, since $L(\phi_n) = -\lambda_n g\phi_n$, ϕ_n is super-(L) when $\phi_n > 0$, and is sub-(L) when $\phi_n < 0$. Moreover, if $L_m = L+\lambda_m g$, ϕ_n is super-(L_m) when $\phi_n > 0$, and is sub-(L_m) when $\phi_n < 0$, for every $m < n$.

I am indebted to Professor W. W. Rogosinski for much helpful criticism.

REFERENCES

1. G. Valiron, 'Fonctions convexes et fonctions entières', *Bull. Soc. Math. France*, 60 (1932), 278–87.
2. E. F. Beckenbach, 'Generalized convex functions', *Bull. American Math. Soc.* 43 (1937), 363–71.
3. E. F. Beckenbach and R. H. Bing, 'On generalized convex functions', *Trans. American Math. Soc.* 58 (1945), 220–30.
4. G. Pólya, 'Untersuchungen über Lücken und Singularitäten von Potenzreihen', *Math. Zeit.* 29 (1929), 549–60.
5. G. Pólya, 'On the mean-value theorem corresponding to a given linear homogeneous differential equation', *Trans. American Math. Soc.* 24 (1922), 312–24.
6. E. C. Titchmarsh, *Theory of functions* (Oxford, 1932).
7. G. H. Hardy, J. E. Littlewood, G. Pólya, *Inequalities* (Cambridge, 1934).
8. M. M. Peixoto, 'Generalized convex functions and second order differential inequalities', *Bull. American Math. Soc.* 55 (1949), 563–72.

THE MAXIMUM TERM OF AN ENTIRE SERIES (IV)

By S. M. SHAH (*Aligarh*)

[Received 8 March 1949]

1. Introduction

LET $f(z) = \sum_0^\infty a_n z^n$ be an entire function of order ρ $(0 \leqslant \rho \leqslant \infty)$ and lower order λ and let $\mu(r)$ denote the maximum term of the series for $|z| = r$, $\nu(r)$ the rank of this term. It is known [(**1**) 30–2; (**2**)] that

$$\log M(r) \sim \log \mu(r) \quad (0 \leqslant \rho < \infty), \tag{1.1}$$

$$\varliminf_{r\to\infty} \frac{\log \mu(r)}{\nu(r)} \leqslant \frac{1}{\rho} \leqslant \frac{1}{\lambda} \leqslant \varlimsup_{r\to\infty} \frac{\log \mu(r)}{\nu(r)} \quad (0 \leqslant \rho \leqslant \infty). \tag{1.2}$$

Hence for functions of infinite order

$$\varliminf_{r\to\infty} \log \mu(r)/\nu(r) = 0.$$

In this note I prove

THEOREM 1. *For every entire function of infinite order*

$$\varliminf_{r\to\infty} \log M(r)/\nu(r) = 0. \tag{1.3}$$

This is a best possible result, for we have

THEOREM 2. *Given any increasing function $\phi(x)$ tending to infinity (however slowly) with x, there exists an entire function of infinite order for which*

$$\lim_{r\to\infty} \frac{\phi(r) \log M(r)}{\nu(r)} = \infty. \tag{1.4}$$

THEOREM 3. *Let $L(x)$ be any logarithmico-exponential function* [(**3**) 17] *tending to infinity with x and let*

$$\begin{matrix}\varlimsup \\ \varliminf\end{matrix}_{r\to\infty} \frac{\log\log \mu(r)}{L(r)} = \begin{cases} M \\ m. \end{cases}$$

Then $$\varliminf_{r\to\infty} \frac{\log \mu(r) r L'(r)}{\nu(r)} \leqslant \frac{1}{M} \leqslant \frac{1}{m} \leqslant \varlimsup_{r\to\infty} \frac{\log \mu(r) r L'(r)}{\nu(r)}. \tag{1.5}$$

COROLLARY 1. *If*

$$\log\log \mu(r) = \{1+o(1)\} \log\log r \tag{1.6}$$

for a sequence of values of r tending to infinity, then

$$\varlimsup_{r\to\infty} \frac{\log \mu(r)}{\nu(r) \log r} = 1. \tag{1.7}$$

COROLLARY 2. *If* $\log\mu(r) \sim k\nu(r)\log r$ $(0 \leqslant k \leqslant 1)$, *then*

$$\lim_{r\to\infty}\frac{\log\log\mu(r)}{\log\log r} = \frac{1}{k}. \tag{1.8}$$

The inequality (1.2) follows from Theorem 3 if we take $L(r) = \log r$. For $f(z)$ of any order, we have

$$\varlimsup_{r\to\infty}\frac{\log\mu(r)}{\nu(r)\log r} \leqslant 1. \tag{1.9}$$

Corollary 1 shows that for a class of functions we have a precise relation (1.7). It is not possible to improve the hypothesis (1.6); for, given any positive constant β, we can construct a function $f(z)$ for which

$$\log\log\mu(r) = \{1+\beta+o(1)\}\log\log r$$

for all $r > r_0$ and for which

$$\lim_{r\to\infty}\frac{\log\mu(r)}{\nu(r)\log r} = \frac{1}{1+\beta} < 1.$$

Take, for instance

$$f(z) = \sum_{1}^{\infty}\frac{z^n}{n^{n^\alpha}}, \qquad \alpha = \frac{1+\beta}{\beta}.$$

2. Proof of Theorem 1

We have [(1) 31]

$$M(r) \leqslant \mu(r)\left(p+\frac{r}{R_p-r}\right),$$

where p is an integer greater than $\nu(r)$ such that $R_p > r$. Take

$$p = \nu\left(r+\frac{1}{r\nu^2(r)}\right)+1.$$

Then
$$R_p > r+1/r\nu^2(r)$$
and

$$M(r) \leqslant \mu(r)\left\{\nu\left(r+\frac{1}{r\nu^2(r)}\right)+1+r^2\nu^2(r)\right\},$$

$$\log M(r) \leqslant \{1+o(1)\}\log\mu(r)+2\log\nu\left(r+\frac{1}{r\nu^2(r)}\right) \tag{2.1}$$

for $r > r_0$. Further, from (1.2),

$$\varliminf_{r\to\infty}\log\mu(r)/\nu(r) = 0.$$

Hence
$$\varliminf_{n\to\infty}\log\mu(R_n)/\nu(R_n) = 0.$$

Given $\epsilon > 0$, let E denote the sequence of all positive integers $n_1, n_2,\ldots$ such that

$$\log\mu(R_m)/\nu(R_m) < \epsilon \quad (m = n_1, n_2,\ldots).$$

Either [Case A] there is a sub-sequence of integers K_t say ($t = 1, 2, 3,...$) tending to infinity such that

$$R_{m+1} > R_m + \frac{1}{m^2} \quad (m = K_t), \tag{2.2}$$

in which case

$$\nu\left(R_m + \frac{1}{R_m \nu^2(R_m)}\right) = \nu(R_m),$$

$$\frac{\log M(R_m)}{\nu(R_m)} < \{1+o(1)\}\frac{\log \mu(R_m)}{\nu(R_m)} + \frac{2\log \nu(R_m)}{\nu(R_m)} < 3\epsilon \text{ for } m > m_0. \tag{2.3}$$

Or [Case B] for all large m, say $m > N$, of E

$$R_{m+1} \leqslant R_m + \frac{1}{m^2}, \tag{2.4}$$

in which case either $R_{m+1} = R_m$ and then $m+1 \subset E$ or $R_{m+1} > R_m$,

$$\frac{\log \mu(R_{m+1})}{\nu(R_{m+1})} \leqslant \frac{1}{m+1}\left\{\log \mu(R_m) + \int\limits_{R_m}^{R_{m+1}} \frac{\nu(x)}{x}\,dx\right\}$$

since $\nu(R_{m+1}) \geqslant m+1$. Now, from (2.4),

$$m \log \frac{R_{m+1}}{R_m} < \frac{1}{mR_m} < \frac{\log \mu(R_m)}{m}.$$

Hence

$$\frac{\log \mu(R_{m+1})}{\nu(R_{m+1})} < \frac{\log \mu(R_m)}{m} < \epsilon,$$

and so $m+1 \subset E$. Similarly $m+2, m+3,... \subset E$. Let $m \subset E$ and $m > N$. Then

$$R_{m+p} \leqslant R_m + \sum_{m}^{m+p-1} \frac{1}{n^2} < \text{a constant},$$

which leads to a contradiction since R_{m+p} tends to infinity with p. Hence the alternative (B) is not possible and (2.3) holds and the theorem is proved.

3. Proof of Theorem 2

Let $\phi(r) > 0$ for $r \geqslant r_0$ and let

$$e_1(x) = \exp(x), \qquad e_{k+1}(x) = \exp\{e_k(x)\}$$
$$l_1(x) = \log x, \qquad l_{k+1}(x) = \log\{l_k(x)\}.$$

We can construct a non-decreasing function $\psi(r)$ tending to infinity with r such that

$$\psi(r) = 0 \quad \text{for} \quad 0 \leqslant r \leqslant r_1 = \max\{r_0, e_4(1)\},$$
$$\psi(r) \leqslant \sqrt{\{\phi(r)\}} \quad \text{for} \quad r \geqslant r_1,$$
$$\psi'(r) \leqslant \frac{1}{rl_1(r)l_2(r)l_3(r)} \quad \text{for} \quad r \geqslant r_1, \tag{3.1}$$

except possibly at an enumerable set of points where the right-hand and the left-hand derivatives $\psi'(r\pm 0)$ exist and satisfy (3.1). Let N be a positive integer such that $\psi(N) \geqslant 1$.

LEMMA 1. *For all* $n \geqslant N$,

$$n^{1/\psi(n)} < (n+1)^{1/\psi(n+1)}.$$

For

$$\psi(n+1)-\psi(n) = \int_n^{n+1} \psi'(t)\,dt \leqslant \int_n^{n+1} \frac{dt}{tl_1(t)l_2(t)l_3(t)}$$

$$\leqslant \frac{1}{l_1(n)l_2(n)l_3(n)}\log\left(1+\frac{1}{n}\right),$$

$$\frac{\psi(n+1)}{\psi(n)} \leqslant 1+\frac{1}{l_1(n)l_2(n)l_3(n)}\frac{1}{\psi(n)}\log\left(1+\frac{1}{n}\right)$$

$$< 1+\frac{1}{\log n}\log\left(1+\frac{1}{n}\right) = \frac{\log(n+1)}{\log n},$$

which proves the lemma.

Let $$R_1 = R_2 = \ldots = R_N = 1, \quad R_m = m^{1/\psi(m)}$$

for $m = N+1,\, N+2,\ldots$. Then, for $n \geqslant N$, R_n is a steadily increasing function of n tending to infinity with n.

LEMMA 2. $$\sum_1^n \log R_m = \frac{n}{\psi(n)}\{\log n - 1 + o(1)\}.$$

For

$$\sum_1^n \log R_m = \sum_{N+1}^n \frac{\log m}{\psi(m)}$$

$$= \int_{N+1}^n \frac{\log x\,dx}{\psi(x)} + O\left\{\frac{\log n}{\psi(n)}\right\}$$

$$= \frac{n\log n}{\psi(n)} - \int_{N+1}^n \frac{dx}{\psi(x)} + \int_{N+1}^n \frac{x\log x\,\psi'(x)\,dx}{\psi^2(x)} + O\left\{\frac{\log n}{\psi(n)}\right\}$$

$$= \frac{n\log n - n + O(\log n)}{\psi(n)} - \int_{N+1}^n \frac{x\,\psi'(x)\,dx}{\psi^2(x)} + \int_{N+1}^n \frac{x\log x\,\psi'(x)\,dx}{\psi^2(x)}.$$

Let the two integrals in the last expression be denoted by I_1 and I_2. Then

$$I_2 < \int_{N+1}^n \frac{x\log x\,dx}{xl_1(x)l_2(x)l_3(x)\psi^2(x)} = \left[\int_{N+1}^{\sqrt{n}} + \int_{\sqrt{n}}^n\right] \leqslant \left[O(\sqrt{n}) + \frac{A}{\psi^2(\sqrt{n})}\frac{n}{l_2(n)l_3(n)}\right].$$

Now $$\psi(n)-\psi(\sqrt{n}) = \int\limits_{\sqrt{n}}^{n} \psi'(t)\,dt < \frac{B}{l_2(n)l_3(n)} \to 0 \quad \text{as} \quad n \to \infty.$$

Hence $I_2 = o\{n/\psi(n)\}$. Similarly $I_1 = o\{n/\psi(n)\}$ and the lemma follows.

Let $$f(z) = \sum_{1}^{\infty} \frac{z^n}{R_1 R_2 \dots R_n}.$$

From Lemma 2 it follows that $f(z)$ is an entire function of infinite order. Further, for $R_n \leqslant r < R_{n+1}$ $(n > n_0 > N)$,

$$\nu(r) = n, \qquad \mu(r) = \frac{r^n}{R_1 R_2 \dots R_n},$$

$$\frac{\log \mu(r)}{\nu(r)} \geqslant \frac{1}{n} \log \mu(R_n) = \frac{1}{n}\Big\{n \log R_n - \sum_{m=1}^{n} \log R_m\Big\} = \frac{1+o(1)}{\psi(n)},$$

$$\frac{\phi(r)\log \mu(r)}{\nu(r)} \geqslant \frac{\phi(R_n)}{\psi(n)}\{1+o(1)\} \geqslant \psi(R_n)\frac{\psi(R_n)}{\psi(n)}\{1+o(1)\},$$

which tends to infinity with n, since $\psi(R_n) \sim \psi(n)$. Hence the theorem is proved.

4. Proof of Theorem 3

The proof is omitted as being similar to that of Lemma 1 of (**4**). The two corollaries follow immediately from the theorem.

REFERENCES

1. G. Valiron, *Lectures on the General Theory of Integral Functions* (Toulouse, 1923).
2. S. M. Shah, 'The maximum term of an entire series', *Math. Student*, 10 (1942), 80–2.
3. G. H. Hardy, *Orders of Infinity* (Cambridge, 1924).
4. S. M. Shah, 'A theorem on integral functions of integral order', *J. of London Math. Soc.* 15 (1940), 23–31.

ON THE STABILITY OF MACLAURIN SPHEROIDS ROTATING WITH CONSTANT ANGULAR VELOCITY

By C. PRASAD (*Oxford*)

[Received 15 March 1949]

DISCUSSING linear series in his writings on the configurations of rotating fluid masses, Jeans (**1**) has written 'it appears that a change of stability occurs at every point of bifurcation, and at every point on a linear series at which μ [the slowly varying parameter of the series] passes through a maximum or a minimum value.'

In the case of a Maclaurin spheroid rotating with constant angular velocity this statement leads to a result which is contrary to what one would suppose intuitively. As the angular velocity increases, the Maclaurin spheroids, which are stable at first, become unstable after the point of bifurcation with the Jacobian ellipsoids is crossed. If the angular velocity is further increased, we get a series of unstable spheroids until the point is reached for which the angular velocity is a maximum. After this point, increasing ellipticity of the Maclaurin spheroids accompanies *decreasing* angular velocity, and the question arises whether the series becomes *stable* after the turning point, as Jeans' statement would imply. This led me to investigate the problem analytically.

I shall start with the general case of an ellipsoidal configuration. Let ω be the angular velocity, W the gravitational potential energy, and I the moment of inertia about the axis of rotation. For the configurations $\omega =$ constant,

$$W-\tfrac{1}{2}\omega^2 I$$

must be stationary for equilibrium, and an absolute minimum for secular stability (**2**).

Now, we have

$$\begin{aligned} W &= -\tfrac{1}{2}\int V\,dm \\ &= -\tfrac{1}{2}\pi\gamma\rho abc \iiint\int_0^\infty \left(1-\frac{x^2}{a^2+\lambda}-\frac{y^2}{b^2+\lambda}-\frac{z^2}{c^2+\lambda}\right)\frac{d\lambda}{\Delta(\lambda)}\rho\,dxdydz, \end{aligned}$$

where the triple integral extends over the volume of the ellipsoid and

$$\Delta(\lambda) = \sqrt{\{(a^2+\lambda)(b^2+\lambda)(c^2+\lambda)\}}.$$

Performing the triple integration, we get after some simplification

$$W = -\tfrac{8}{15}\gamma(\pi\rho abc)^2 \int\limits_0^\infty \frac{d\lambda}{\Delta(\lambda)},$$

and putting

$$\lambda = c^2x, \qquad s = c^2/a^2, \qquad t = c^2/b^2, \qquad \tfrac{4}{3}\pi\rho abc = \tfrac{4}{3}\pi\rho r_0^3 = M,$$

we may write
$$W = -\tfrac{2}{5}\pi\gamma\rho M r_0^2 (st)^{\frac{1}{3}} \int\limits_0^\infty \frac{dx}{\Delta},$$

where Δ now stands for $\sqrt{\{(1+sx)(1+tx)(1+x)\}}$.

We also have

$$I = \tfrac{1}{5}M(a^2+b^2) = \tfrac{1}{5}Mr_0^2(st)^{\frac{1}{3}}\left(\frac{s+t}{st}\right).$$

Thus
$$W-\tfrac{1}{2}\omega^2 I = -\tfrac{1}{10}Mr_0^2\left[4\pi\gamma\rho(st)^{\frac{1}{3}} \int\limits_0^\infty \frac{dx}{\Delta} + (st)^{\frac{1}{3}}\left(\frac{s+t}{st}\right)\omega^2\right].$$

This gives $W-\frac{1}{2}\omega^2 I$ as a function of the two parameters s and t. Dividing throughout by $(\frac{1}{10}Mr_0^2)(2\pi\gamma\rho)$ and for brevity writing h and W' for $\omega^2/2\pi\gamma\rho$ and $(W-\frac{1}{2}\omega^2 I)/\{(\frac{1}{10}Mr_0^2)(2\pi\gamma\rho)\}$, we have the relation

$$W' = -(st)^{\frac{1}{3}}\left[2\int\limits_0^\infty \frac{dx}{\Delta} + \left(\frac{s+t}{st}\right)h\right]. \tag{1}$$

The stationary values of W' give the configurations of equilibrium, and they will be stable if W' is an absolute minimum. On differentiating (1) and simplifying we get

$$\left.\begin{aligned} \frac{\partial W'}{\partial s} &= -\frac{(st)^{\frac{1}{3}}}{3s}\left[3\int\limits_0^\infty \frac{dx}{(1+sx)\Delta} - \int\limits_0^\infty \frac{dx}{\Delta} + \frac{(s-2t)h}{st}\right], \\ \frac{\partial W'}{\partial t} &= -\frac{(st)^{\frac{1}{3}}}{3t}\left[3\int\limits_0^\infty \frac{dx}{(1+tx)\Delta} - \int\limits_0^\infty \frac{dx}{\Delta} + \frac{(t-2s)h}{st}\right]. \end{aligned}\right\} \tag{2}$$

These equated to zero give two equations, from which by subtraction and by elimination of h we get either $s = t$, in which case the equilibrium configurations are Maclaurin spheroids; or

$$h = st\int\limits_0^\infty \frac{x\,dx}{(1+sx)(1+tx)\Delta} \tag{3}$$

and
$$\int\limits_0^\infty \frac{3+(s+t)x}{(1+sx)(1+tx)\Delta}\,dx - \int\limits_0^\infty \frac{dx}{\Delta} = 0, \tag{4}$$

which give the equilibrium configurations known as Jacobian ellipsoids.

I proceed to find the stability of Maclaurin spheroids. We have on differentiating (2) and simplifying

$$\left.\begin{aligned}
\frac{\partial^2 W'}{\partial s^2} &= -\frac{(st)^{\frac{1}{3}}}{18s^2}\left[27\int_0^\infty \frac{dx}{(1+sx)^2\Delta}-42\int_0^\infty \frac{dx}{(1+sx)\Delta}+\right.\\
&\qquad\qquad\left.+7\int_0^\infty \frac{dx}{\Delta}+4\left(\frac{5t-s}{st}\right)h\right]\\
\frac{\partial^2 W'}{\partial s\partial t} &= -\frac{(st)^{-\frac{2}{3}}}{18}\left[9\int_0^\infty \frac{dx}{(1+sx)(1+tx)\Delta}-3\int_0^\infty \frac{dx}{(1+sx)\Delta}-\right.\\
&\qquad\qquad\left.-3\int_0^\infty \frac{dx}{(1+tx)\Delta}+\int_0^\infty \frac{dx}{\Delta}-4\left(\frac{s+t}{st}\right)h\right]\\
\frac{\partial^2 W'}{\partial t^2} &= -\frac{(st)^{\frac{1}{3}}}{18t^2}\left[27\int_0^\infty \frac{dx}{(1+tx)^2\Delta}-42\int_0^\infty \frac{dx}{(1+tx)\Delta}+\right.\\
&\qquad\qquad\left.+7\int_0^\infty \frac{dx}{\Delta}+4\left(\frac{5s-t}{st}\right)h\right]
\end{aligned}\right\}. \tag{5}$$

For a configuration of equilibrium

$$\delta W' = \frac{1}{2}\left\{\frac{\partial^2 W'}{\partial s^2}(\delta s)^2+2\frac{\partial^2 W'}{\partial s\partial t}(\delta s)(\delta t)+\frac{\partial^2 W'}{\partial t^2}(\delta t)^2\right\},$$

on neglecting higher-order terms. For the particular case of Maclaurin spheroids $s = t$, and this gives

$$\delta W' = \tfrac{1}{2}(\alpha\,\delta s^2+2\beta\,\delta s\delta t+\alpha\,\delta t^2), \tag{6}$$

where $$\alpha = \left(\frac{\partial^2 W'}{\partial s^2}\right)_{s=t} = \left(\frac{\partial^2 W'}{\partial t^2}\right)_{s=t} \quad\text{and}\quad \beta = \left(\frac{\partial^2 W'}{\partial s\partial t}\right)_{s=t}. \tag{7}$$

This may be written as

$$\delta W' = (\alpha+\beta)\left(\frac{\delta s+\delta t}{2}\right)^2+(\alpha-\beta)\left(\frac{\delta s-\delta t}{2}\right)^2,$$

i.e. $$\delta W' = b_1\phi_1^2+b_2\phi_2^2, \tag{8}$$

where

$$b_1 = \alpha+\beta,\quad b_2 = \alpha-\beta,\quad \phi_1 = \tfrac{1}{2}(\delta s+\delta t),\quad \phi_2 = \tfrac{1}{2}(\delta s-\delta t). \tag{9}$$

For stability W' must be an absolute minimum, and therefore b_1 and b_2 must both be positive.

We have from (5), (7), (9)

$$b_1 = -\tfrac{2}{9}s^{-\frac{4}{3}}\left[9\int\limits_0^\infty \frac{dx}{(1+sx)^2\Delta} - 12\int\limits_0^\infty \frac{dx}{(1+sx)\Delta} + 2\int\limits_0^\infty \frac{dx}{\Delta} + \frac{2}{s}h\right] \tag{10}$$

where now $\Delta = (1+sx)\surd(1+x)$.

For the case $s = t$, we have from the equation $\partial W'/\partial s = 0$,

$$h = 3s\int\limits_0^\infty \frac{dx}{(1+sx)\Delta} - s\int\limits_0^\infty \frac{dx}{\Delta}. \tag{11}$$

From this we get

$$\frac{dh}{ds} = 6\int\limits_0^\infty \frac{dx}{(1+sx)^2\Delta} - 4\int\limits_0^\infty \frac{dx}{(1+sx)\Delta}.$$

Now, substituting the value of h from (11) in (10), we get

$$b_1 = -\tfrac{2}{9}s^{-\frac{4}{3}}\left[9\int\limits_0^\infty \frac{dx}{(1+sx)^2\Delta} - 6\int\limits_0^\infty \frac{dx}{(1+sx)\Delta}\right] = -\tfrac{1}{3}s^{-\frac{4}{3}}\frac{dh}{ds}$$

$$= \frac{(1-e^2)^{-\frac{4}{3}}}{6e}\frac{dh}{de}$$

since $s = c^2/a^2 = 1-e^2$.

We know [**3**] that h increases with e at first, attains a maximum at $e = 0{\cdot}930$, and then decreases as e further increases. Hence b_1 is positive when $0 \leqslant e < 0{\cdot}93$, zero at $e = 0{\cdot}93$, and negative when $e > 0{\cdot}93$.

Again, from (5),

$$b_2 = -\tfrac{1}{3}s^{-\frac{4}{3}}\left[3\int\limits_0^\infty \frac{dx}{(1+sx)^2\Delta} - 6\int\limits_0^\infty \frac{dx}{(1+sx)\Delta} + \int\limits_0^\infty \frac{dx}{\Delta} + \frac{4}{s}h\right]$$

$$= -s^{-\frac{4}{3}}\left[\int\limits_0^\infty \frac{dx}{(1+sx)^2\Delta} + 2\int\limits_0^\infty \frac{dx}{(1+sx)\Delta} - \int\limits_0^\infty \frac{dx}{\Delta}\right], \quad \text{by (11)}.$$

Now, if we put $s = t$ in the left-hand side of equation (4), it becomes

$$\int\limits_0^\infty \frac{3+2sx}{(1+sx)^2\Delta}\,dx - \int\limits_0^\infty \frac{dx}{\Delta} = \int\limits_0^\infty \frac{dx}{(1+sx)^2\Delta} + 2\int\limits_0^\infty \frac{dx}{(1+sx)\Delta} - \int\limits_0^\infty \frac{dx}{\Delta} = -s^{\frac{4}{3}}b_2.$$

Thus b_2 is zero for the Maclaurin spheroid which satisfies equation (4), or in other words, for the Maclaurin spheroid at the point of bifurcation, for which $e = 0{\cdot}8127$.

Putting b_2 in terms of l, where

$$s = 1-e^2 = 1/(1+l^2),$$

we get by actual integration

$$b_2 = \frac{3+14l^2+3l^4}{4l^5(1+l^2)^{\frac{1}{2}}}\left[\frac{3l+13l^3}{3+14l^2+3l^4}-\tan^{-1}l\right].$$

The sign of b_2 depends on the terms within square brackets, and, if we put

$$\phi(l) = \frac{3l+13l^3}{3+14l^2+3l^4}-\tan^{-1}l,$$

we have

$$\phi'(l) = \frac{16l^4(3l^2+1)(1-l^2)}{(1+l^2)(3+14l^2+3l^4)^2}.$$

Thus $\phi'(l)$ is positive when $0 < l < 1$ and negative when $l > 1$. Therefore $\phi(l)$ continuously increases in the range $0 < l < 1$ and then continuously decreases when $l > 1$. But, when $l \sim 0$, $\phi(l) \sim \frac{16}{45}l^5$ and, when $l \sim \infty$, $\phi(l) \sim -\frac{1}{2}\pi$. Hence $\phi(l)$, and therefore b_2, remains positive from $l = 0$ to some value l_0 of l, when it becomes zero, and then it remains negative for all values of $l > l_0$. The value of the eccentricity corresponding to l_0 is 0·8127. So we see that b_2 is positive when $e < 0{\cdot}8127$ and negative when $e > 0{\cdot}8127$.

Summarizing the above results, we see that, if $e < 0{\cdot}8127$, both b_1 and b_2 are positive and the Maclaurin spheroids are stable for all kinds of displacements. After the point of bifurcation with Jacobian ellipsoids is crossed, b_2 becomes negative, though b_1 remains positive. So the Maclaurin spheroids are unstable in this case. If, however, they are constrained to remain spheroidal, $\delta s = \delta t$, and thus $\delta W' = b_1\phi_1^2$, which is positive; so they remain stable. For configurations of eccentricity greater than 0·930 (which corresponds to the configuration of maximum angular velocity) b_1 and b_2 are both negative and the spheroids are totally unstable.

Thus it is not in general true that a change of stability always accompanies the passage of the varying parameter (in this case ω) through a maximum or a minimum.

I am grateful to Professor E. A. Milne for his encouragement and interest in the preparation of this paper.

REFERENCES

1. J. H. Jeans, *Astronomy and Cosmogony* (1929), 194.
2. —— ibid. 195–9.
3. F. Tisserand, *Traité de Mécanique Céleste*, 2 (1891), 84.

THE APPLICATION OF MULTIPLE FOURIER TRANSFORMS TO THE SOLUTION OF PARTIAL DIFFERENTIAL EQUATIONS

By J. L. B. COOPER (*London*)

[Received 19 March 1949]

1. Introduction

Although the literature contains a number of applications of multiple Fourier Transforms to the solution of partial differential equations, and in particular to the solution of the initial-value problem of the wave equation [see for example **1** (387–95), **2**, **3**] these solutions are not investigated rigorously. Thus Courant uses Fourier integrals as a heuristic process to obtain a formal solution, which is later justified by other means. The methods used so far can be made rigorous only if the solutions are assumed *a priori* to satisfy certain rather arbitrary conditions. In view of the wide scope of the method for the solution of equations with constant coefficients (at least) it seems worth while to show that it can be used rigorously with no more assumptions than are necessary in any other method of solution. This is the aim of the present paper. The method will be illustrated by application to the wave equation in an arbitrary number of dimensions.

To begin with, the techniques to be used here and the difficulties in the application of multiple Fourier transforms which they are intended to overcome will be explained.

1.1. The first difficulty in any application of Fourier transforms concerns the necessity of some assumption about the behaviour of the function at infinity. If the Fourier transforms are taken over the whole infinite region of integration, a restriction on the size of the functions involved at infinity is unavoidable. I shall use instead a 'truncated' Fourier transform, as in (**4**); this involves boundary terms in the solution, but in the case of hyperbolic differential equations these can be eliminated. (Parabolic and elliptic differential equations need further examination: their solution, or proof of uniqueness, generally require restrictions on the size of the solution at infinity, but these are demanded by the equation rather than by the method.) One advantage of this method is that the uniqueness problem of the differential equation is solved; whereas the ordinary Fourier-transform method proves uniqueness only within the class of solutions satisfying the conditions at infinity imposed by its use.

1.2. The second difficulty is peculiar to multiple, as opposed to simple, Fourier transforms. Convergence or summability of an ordinary Fourier integral depends only on the local properties of the transformed function; convergence or summability of multiple Fourier transforms, by most methods, depends on properties of the function near any point which has one coordinate the same as that of the point at which convergence is being considered. The effect of this is to make convergence very difficult to prove for the integrals obtained as solutions of differential equations. For this reason I shall use here the spherical means of S. Bochner, for which convergence is a local property [see **5**]. We need consider for our purpose only spherical Abelian means: that is to say, we evaluate the multiple Fourier integral for $f(x_1, x_2,..., x_k)$ by considering

$$\lim_{\delta\to 0}\frac{1}{(2\pi)^k}\int\int e^{-\delta\rho-iu.x}\,du_1...du_k\int\int f(y_1,..., y_k)e^{iu.y}\,dy_1...dy_k, \tag{1}$$

where
$$\rho^2 = u_1^2+...+u_k^2, \qquad u.x = u_1x_1+u_2x_2+...+u_kx_k. \tag{2}$$

In the second section of this paper preliminary results on the convergence of spherical Abelian means are given; this is done for completeness, and because a much simpler proof can be given in this particular case than is given by Bochner for the very general case which he considers. In the third section, the wave equation in k dimensions is discussed; the solution is carried out in detail for k odd and also for $k = 2$, in the case when the Poisson form of the solution holds. A discussion of the existence of the solution, and the propagation of the solution according to Huygens's principle is then given; this applies even when the conditions of differentiability necessary for the Poisson solution do not hold. This section of the paper proves existence of the solution in a mean-square sense, and does not involve the theory of summability by spherical means: its conclusions are valid even with discontinuous initial conditions.

2. Spherical means

For shortness I shall use the notations

$$du \quad \text{for} \quad du_1\,du_2\,du_3...du_k, \text{ etc.}, \qquad f(u) \quad \text{for} \quad f(u_1, u_2, u_3,..., u_k), \text{ etc.},$$
$$u.x = u_1x_1+u_2x_2+...+u_kx_k,$$
$$\rho^2 = u_1^2+u_2^2+...+u_k^2, \qquad r^2 = x_1^2+x_2^2+...+x_k^2,$$

Integrals with no limits are k-dimensional integrals over the entire space. Integral signs will not be repeated in multiple integrals.

The integral in (1) is of the form

$$\frac{1}{(2\pi)^k}\int \sigma(\rho)\,du \int f(y)e^{iu.(y-x)}\,dy = I_\sigma(x), \text{ say,} \tag{3}$$

with $\sigma(\rho) = e^{-\delta\rho}$. We shall consider integrals of this form, supposing always that $f(x)$ is summable over the entire range of integration. In this case we are justified in inverting the order of integration, and we get

$$I_\sigma(x) = \frac{1}{(2\pi)^k}\int f(y)\,dy \int \sigma(\rho)e^{iu.(y-x)}\,du. \tag{4}$$

The inner integral is a function of r only, where

$$r^2 = (y_1-x_1)^2+...+(y_k-x_k)^2,$$

and it can be written [see **6**, 186]

$$H_\sigma(r) = \int \sigma(\rho)e^{iu.(y-x)}\,du = \frac{(2\pi)^{\frac{1}{2}k}}{r^{\frac{1}{2}(k-2)}}\int\limits_0^\infty \sigma(\rho)\rho^{\frac{1}{2}k}J_{\frac{1}{2}(k-2)}(\rho r)\,d\rho. \tag{5}$$

$I_\sigma(x)$ can now be evaluated by summation by spherical shells, and we get

$$I_\sigma(x) = \frac{\omega_k}{(2\pi)^k}\int\limits_0^\infty f_x(r)r^{k-1}H_\sigma(r)\,dr, \tag{6}$$

where I have introduced the following notation: for any function $\psi(x)$, $\psi_x(r)$ denotes the mean value of ψ over a sphere of radius r about the point x: that is to say,

$$\psi_x(r) = \omega_k^{-1}\int\limits_{S_k} \psi(x_1+r\xi_1,..., x_k+r\xi_k)\,dS_k,$$

where the integral is taken over the surface of the k-dimensional unit sphere S_k ($\xi_1^2+...+\xi_k^2 = 1$) and

$$\omega_k = 2\pi^{\frac{1}{2}k}/\Gamma(\tfrac{1}{2}k)$$

is the surface area of this sphere.

Spherical means. Consider the integral (1), which is the special case of (3) with $\sigma(\rho) = e^{-\delta\rho}$. In this case

$$H_\sigma(r) = (2\pi)^{\frac{1}{2}k}r^{-\frac{1}{2}(k-2)}\int\limits_0^\infty e^{-\delta\rho}\rho^{\frac{1}{2}k}J_{\frac{1}{2}(k-2)}(r\rho)\,d\rho$$

$$= 2^k\pi^{\frac{1}{2}(k-1)}\Gamma(\tfrac{1}{2}k+\tfrac{1}{2})\delta/(\delta^2+r^2)^{\frac{1}{2}(k+1)},$$

from known Bessel-function results. Hence

$$I_\sigma(x) = \frac{2\Gamma[\frac{1}{2}(k+1)]}{\sqrt{\pi}\,\Gamma(\frac{1}{2}k)} \int\limits_0^\infty f_x(r) \frac{\delta r^{k-1}\,dr}{(\delta^2+r^2)^{\frac{1}{2}k+\frac{1}{2}}}.$$

We must consider the limit of this expression as $\delta \to 0$. For all $\delta > 0$,

$$\int\limits_0^\infty \frac{\delta r^{k-1}\,dr}{(\delta^2+r^2)^{\frac{1}{2}(k+1)}} = \frac{\Gamma(\frac{1}{2})\Gamma(\frac{1}{2}k)}{2\Gamma(\frac{1}{2}k+\frac{1}{2})},$$

so that, for any constant s,

$$|I_\sigma(x)-s| = \frac{2\Gamma(\frac{1}{2}k+\frac{1}{2})}{\sqrt{\pi}\,\Gamma(\frac{1}{2}k)} \int\limits_0^\infty |f_x(r)-s| \frac{\delta r^{k-1}\,dr}{(\delta^2+r^2)^{\frac{1}{2}k+\frac{1}{2}}}. \tag{7}$$

Write $$F(t) = \int\limits_0^t |f_x(r)-s| r^{k-1}\,dr.$$

I shall now prove that (7) tends to zero as $\delta \to 0$ if $F(t)t^{-k} \to 0$ as $t \to 0$. For any $\epsilon > 0$, there then exists $\eta > 0$ such that $F(t) < \epsilon t^k$ for $t < \epsilon$. Take $\delta < \eta$ and split the integral in (7) into contributions from $(0,\delta)$ and (δ,∞). Then the integral over (δ,∞) gives, after integration by parts,

$$[\delta F(r)(\delta^2+r^2)^{-\frac{1}{2}k-\frac{1}{2}}]_\delta^\infty + (k+1)\int\limits_\delta^\infty \delta r\, F(r)(\delta^2+r^2)^{-\frac{1}{2}k-\frac{3}{2}}\,dr.$$

The expression in brackets is $2^{-k-1}\delta^{-k}F(\delta)$ and this tends to zero as $\delta \to 0$. As for the integral, consider the contributions to it from (δ,η) and (η,∞). $F(t)$ is $O(t^k)$ as $t \to \infty$, since $f_x(r)r^{k-1}$ is $L(0,\infty)$. Hence the integral over (η,∞) tends to zero as $\delta \to 0$. The contribution from (δ,η) is less than

$$\epsilon\delta \int\limits_\delta^\eta r^{-2}\,dr = \epsilon\delta[\delta^{-1}-\eta^{-1}] < \epsilon.$$

As for the contribution from $(0,\delta)$ to the integral in (7), we have, for $0 \leqslant r \leqslant \delta$,

$$\delta r^{k-1}(\delta^2+r^2)^{-\frac{1}{2}k-\frac{1}{2}} \leqslant \delta^{-k} r^{k-1},$$

so that the contribution of this part of the integral is less than

$$\delta^{-k}\int\limits_0^\delta |f_x(r)-s| r^{k-1}\,dr = \delta^{-k}F(\delta).$$

Hence the integral $I_\sigma(x)$ tends to s if

$$\int\limits_0^t |f_x(r)-s|r^{k-1}\,dr = o(t^k) \quad \text{as} \quad t \to 0. \tag{8}$$

If we take $s = f(x)$, then

$$|f_x(t)-s| \leqslant \frac{1}{\omega_k}\int\limits_{S_k} |f(x_1+t\xi_1,..., x_k+t\xi_k)-f(x_1,..., x_k)|\, dS_k \to 0$$

as $t \to 0$, if $f(x)$ is continuous at x. Thus, for any continuous function, and, more generally, for any summable function for which (8) is satisfied with $s = f(x)$,

$$f(x) = \lim_{\delta\to 0}\frac{1}{(2\pi)^k}\int e^{-\delta\rho}\,du \int f(y)e^{iu.(y-x)}\,dy. \tag{9}$$

3. The wave equation

Consider now the k-dimensional wave equation

$$\nabla^2 f = \frac{\partial^2 f}{\partial x_1^2}+...+\frac{\partial^2 f}{\partial x_k^2} = \frac{\partial^2 f}{\partial t^2}, \tag{10}$$

with the initial conditions $f(x,0) = \psi(x)$, $f_t(x,0) = \phi(x)$.

To begin with, we restrict ourselves to finding the solution for values of t less in absolute value than some constant T, and values of x in some bounded region D'. Let D be a domain, which we may take to be a sphere or a cube, with surface S, lying in the x-space and such that a sphere of radius $2T$ about any point of D' lies wholly inside D. Write

$$F_D(u,t) = \int\limits_D f(x,t)e^{iu.x}\,dx,$$

and, correspondingly,

$$\Phi_D(u) = \int\limits_D \phi(x)e^{iu.x}\,dx, \qquad \Psi_D(u) = \int\limits_D \psi(x)e^{iu.x}\,dx.$$

By Green's theorem

$$\int\limits_D [\nabla^2 f e^{iu.x}+\rho^2 f e^{iu.x}]\,dx = \int\limits_S \left(\frac{\partial f}{\partial \nu}e^{iu.x}-f\frac{\partial}{\partial \nu}e^{iu.x}\right)dS,$$

and so the wave equation (10) gives

$$\frac{\partial^2 F_D}{\partial t^2}+\rho^2 F_D(u,t) = \chi_D(u,t), \tag{11}$$

where χ_D denotes the surface integral on the right-hand side of the previous equation. We get, solving (11),

$$F_D(u,t) = A(u)\sin\rho t + B(u)\cos\rho t + \frac{1}{\rho}\int\limits_0^t \chi_D(u,\tau)\sin\rho(t-\tau)\,d\tau,$$

and, from the initial conditions,

$$\rho A(u) = \Phi_D(u), \qquad B(u) = \Psi_D(u).$$

Since $f(x,t)$ is continuous in x, it follows from the last section that, for all x inside D,

$$f(x,t) = \lim_{\delta\to 0}\frac{1}{(2\pi)^k}\int e^{-\delta\rho+iu.x}F_D(u,t)\,du,$$

and, on substituting the expression found for $F_D(u,t)$, we get

$$f(x,t) = \lim_{\delta\to 0}\frac{1}{(2\pi)^k}\left\{\int e^{-\delta\rho-iu.x}\frac{\sin\rho t}{\rho}\,du\int\limits_D \phi(y)e^{iu.y}\,dy + \right.$$

$$+\int e^{-\delta\rho-iu.x}\cos\rho t\,du\int\limits_D \psi(y)e^{iu.y}\,dy +$$

$$\left. + \int e^{-\delta\rho-iu.x}\frac{du}{\rho}\int\limits_0^t \chi_D(u,\tau)\sin\rho(t-\tau)\,d\tau\right\}. \quad (12)$$

This is an expression for the solution of the wave equation, which can be simplified to give the classical expressions. I shall now carry out this simplification, supposing that *k is odd.*

The first term in (12) is a case of the integral (3) with ϕ in place of f, and

$$\sigma(\rho) = e^{-\delta\rho}\frac{\sin\rho t}{\rho}.$$

In this case

$$H_\sigma(r) = (2\pi)^{\frac{1}{2}k}r^{-\frac{1}{2}(k-2)}\int\limits_0^\infty e^{-\delta\rho}\sin\rho t\,\rho^{\frac{1}{2}k-1}J_{\frac{1}{2}k-1}(r\rho)\,d\rho$$

$$= (-1)^{\frac{1}{2}(k-1)}(2\pi)^{\frac{1}{2}k}r^{-\frac{1}{2}(k-2)}\left(\frac{\partial}{\partial t}\right)^{k-2}\int\limits_0^\infty e^{-\delta\rho}\cos\rho t\,\rho^{-\frac{1}{2}(k-2)}J_{\frac{1}{2}(k-2)}(r\rho)\,d\rho$$

$$= \frac{(-1)^{\frac{1}{2}(k-1)}2\pi^{\frac{1}{2}k}}{\sqrt{\pi}\,\Gamma(\frac{1}{2}k-\frac{1}{2})}\left(\frac{\partial}{\partial t}\right)^{k-2}\int\limits_0^1 (1-v^2)^{\frac{1}{2}(k-3)}\left\{\frac{\delta}{\delta^2+(t-rv)^2}+\frac{\delta}{\delta^2+(t+rv)^2}\right\}dv.$$

using the known result for the cosine transform of a Bessel function [7, 178]. This can be written

$$H_\sigma(r) = \frac{2(-\pi)^{\frac{1}{2}k-\frac{1}{2}}}{r^{k-2}\Gamma(\frac{1}{2}k-\frac{1}{2})}\left(\frac{\partial}{\partial t}\right)^{k-2} \int\limits_0^r (r^2-v^2)^{\frac{1}{2}(k-3)}[K(\delta,t-v)+K(\delta,t+v)]\,dv,$$

where $$K(\delta,x) = \delta/(\delta^2+x^2).$$

Since the integral can be differentiated under the sign of integration, it can be written

$$\int\limits_0^r (r^2-v^2)^{\frac{1}{2}(k-3)}\left(\frac{\partial}{\partial v}\right)^{k-2}\{K(\delta,t+v)-K(\delta,t-v)\}\,dv.$$

Substituting this in (3), we get

$$I_\sigma(x)$$

$$= \frac{(-1)^{\frac{1}{2}k-\frac{1}{2}}}{\pi\Gamma(k-1)}\int\limits_0^\infty r\phi_x(r)\,dr \int\limits_0^r (r^2-v^2)^{\frac{1}{2}k-\frac{3}{2}}\left(\frac{\partial}{\partial v}\right)^{k-2}[K(\delta,t+v)-K(\delta,t-v)]\,dv$$

$$= \frac{(-1)^{\frac{1}{2}k-\frac{1}{2}}}{\pi\Gamma(k-1)}\int\limits_0^\infty \left(\frac{\partial}{\partial v}\right)^{k-2}[K(\delta,t+v)-K(\delta,t-v)]\,dv \int\limits_v^\infty (r^2-v^2)^{\frac{1}{2}k-\frac{3}{2}}r\phi_x(r)\,dr.$$

After repeated integrations by parts, the integral becomes

$$\left[\sum_{r=0}^{k-3}(-1)^r\frac{\partial^r Z}{\partial v^r}\frac{\partial^{k-3-r}}{\partial v^{k-3-r}}\{K(\delta,t+v)-K(\delta,t-v)\}\right]_0^\infty +$$

$$+(-1)^{k-2}\int\limits_0^\infty [K(\delta,t+v)-K(\delta,t-v)]\frac{\partial^{k-2}Z}{\partial v^{k-2}}\,dv,$$

where we write $$Z(v) = \int\limits_v^\infty (r^2-v^2)^{\frac{1}{2}k-\frac{3}{2}}r\phi_x(r)\,dr.$$

As $\delta \to 0$ the expression in square brackets is $O(\delta)$. For $t > 0$, the integral tends to

$$-\pi\frac{\partial^{k-2}}{\partial t^{k-2}}\int\limits_t^\infty (r^2-t^2)^{\frac{1}{2}k-\frac{3}{2}}r\phi_x(r)\,dr$$

if this expression is a continuous function of t, or if t is in the Lebesgue set of this function [cf. 7, 31]. Since

$$\int\limits_0^\infty r\phi_x(r)(r^2-t^2)^{\frac{1}{2}(k-3)}\,dr$$

is a polynomial of degree $k-3$ in t, its derivative of order $k-2$ vanishes, and hence the expression above is equal to

$$\pi \frac{\partial^{k-2}}{\partial t^{k-2}} \int\limits_0^t (r^2-t^2)^{\frac{1}{2}(k-3)} r\phi_x(r)\, dr.$$

Finally, we get

$$\lim_{\delta\to 0}(2\pi)^{-k} \int e^{-\delta\rho} \frac{\sin \rho t}{\rho} e^{-iu.x}\, du \int \phi(y) e^{iu.y}\, dy$$

$$= \frac{1}{(k-2)!}\left(\frac{\partial}{\partial t}\right)^{k-2} \int\limits_0^t (t^2-r^2)^{\frac{1}{2}(k-3)} r\phi_x(r)\, dr, \quad (13)$$

provided that the last expression is continuous in t; and in any case the statement is true almost everywhere if the expression exists everywhere. If $k = (2m+1)$, it is easy to show by induction that (13) is equal to

$$\sum_{\nu=0}^{m} a_\nu t^\nu \left(\frac{\partial}{\partial t}\right)^{\nu} \{t\phi_x(t)\}, \quad (14)$$

where the a_ν are constants.

The existence of the limit in the form given thus depends on the existence and continuity of the derivatives of $\phi_x(t)$ with respect to t up to the mth order. From the expression for $\phi_x(t)$, these conditions will hold if $\phi(x_1,..., x_k)$ is differentiable with respect to $x_1,..., x_k$ up to the mth order.

The second term in (12) is the value of (3) when

$$\sigma(\rho) = e^{-\delta\rho} \cos \rho t.$$

Here

$$H_\sigma(r) = (2\pi)^{\frac{1}{2}k} r^{-\frac{1}{2}(k-2)} \int\limits_0^\infty e^{-\delta\rho} \cos \rho t\, \rho^{\frac{1}{2}k} J_{\frac{1}{2}(k-2)}(r\rho)\, d\rho$$

$$= (-1)^{\frac{1}{2}(k-1)} (2\pi)^{\frac{1}{2}k} r^{-\frac{1}{2}(k-2)} \left(\frac{\partial}{\partial t}\right)^{k-1} \int\limits_0^\infty e^{-\delta\rho} \cos \rho t\, \rho^{-\frac{1}{2}(k-2)} J_{\frac{1}{2}(k-2)}(r\rho)\, d\rho,$$

and by arguments similar to those above we find that the limit is

$$\frac{1}{(k-2)!}\left(\frac{\partial}{\partial t}\right)^{k-1} \int\limits_0^t (t^2-r^2)^{\frac{1}{2}(k-3)} r\psi_x(r)\, dr, \quad (15)$$

almost everywhere if $\psi(x_1,..., x_k)$ has derivatives up to the $(m+1)$th order, and everywhere if the derivative of order $m+1$ is continuous.

We now consider the last term in (12). It is

$$\int e^{-\delta\rho-iu.x}\frac{du}{\rho}\int\limits_0^t \sin\rho(t-\tau)\,d\tau\int\limits_S\left[\frac{\partial f}{\partial\nu}e^{iu.y}-f\frac{\partial}{\partial\nu}e^{iu.y}\right]dS$$

$$=\int\limits_0^t d\tau\int\limits_S\frac{\partial f}{\partial\nu}\,dS\int e^{-\delta\rho+iu.(y-x)}\frac{\sin\rho(t-\tau)}{\rho}\,du-$$

$$-\int\limits_0^t d\tau\int f\,dS\int e^{-\delta\rho}\frac{\partial}{\partial\nu}e^{iu.(y-x)}\frac{\sin\rho(t-\tau)}{\rho}\,du.$$

The inner integral in the first expression in the bracket has been evaluated above: it is

$$C_k\left(\frac{\partial}{\partial t}\right)^{k-2}r^{-k+2}\int\limits_0^r(r^2-v^2)^{\frac{1}{2}(k-3)}\left\{\frac{\delta}{\delta^2+(t-v)^2}+\frac{\delta}{\delta^2+(t+v)^2}\right\}dv, \tag{16}$$

where C_k is a definite constant.

At this stage, we recall the provision made earlier that we are to consider values of x inside a region D' and values of t such that $r>2t$ everywhere on the boundary of D for any point in D'. If we differentiate the expression (16) any number of times under the integral sign, we see that the integrals so formed converge uniformly to their limits for $0\leqslant t\leqslant T$ and all distances r between points of D' and the boundary of D. The limit of the integral in (16) as $\delta\to 0$ is a multiple of

$$(r^2-t^2)^{\frac{1}{2}(k-3)}$$

and from this uniform convergence we conclude that the inner integral converges uniformly to zero in the range of values of r and t considered. By a similar argument, the second integral above also tends to zero, and we therefore have

$$f(x,t)=\frac{1}{(k-2)!}\left\{\left(\frac{\partial}{\partial t}\right)^{k-2}\int\limits_0^t(t^2-r^2)^{\frac{1}{2}(k-3)}r\phi_x(r)\,dr+\right.$$

$$\left.+\left(\frac{\partial}{\partial t}\right)^{k-1}\int\limits_0^t(t^2-r^2)^{\frac{1}{2}(k-3)}r\psi_x(r)\,dr\right\} \tag{17}$$

almost everywhere in D' if $\phi(x)$ has derivatives up to the order $\frac{1}{2}(k-1)$, and $\psi(x)$ up to order $\frac{1}{2}(k+1)$, and everywhere in D' if the highest of these derivatives are continuous.

The restriction of x to D' and the restriction $t\leqslant T$ is removed simply by noting that D, D', T are quite arbitrary, and are not

involved in the formula (17); so that, for any given point and any given t, the formula (17) can be shown to hold by choosing D' to contain x and $T > t$.

The classical solution of the wave equation is thus found, under the classical conditions, which are known to be necessary for the existence of this form of the solution [cf. **8**, 182, and **1**, 399]. Hadamard demonstrates that, if these conditions are violated, no solution of the Cauchy initial-value problem exists.

So far it has only been shown here that, if the solution exists, it has the form (17). We must also verify that (17) is a solution. To verify that the initial conditions are satisfied, we may use Parseval's theorem

$$\int_D |f(x,t)-\psi(x)|^2\,dx = \frac{1}{(2\pi)^{\frac{1}{2}k}}\int |F_D(u,t)-\Psi_D(u)|^2\,du$$

and, on substituting the expression found for $F_D(u,t)$, it is easily verified that $f(x,t) \to \psi(x)$ in mean square as $t \to 0$, throughout any finite region. The expression (14), with the similar expression involving ψ, then shows that $f(x,t)$ tends to a limit as $t \to 0$; and continuity shows that this limit must be $\psi(x)$ everywhere and not merely almost everywhere.

To show that the equation itself is verified, we observe that, if f_D denotes the function which is equal to f in D and zero outside, then the transforms of $\partial^2 f_D/\partial t^2$ and $\nabla^2 f_D$ are

$$(2\pi)^{-\frac{1}{2}k}\frac{\partial^2 F_D}{\partial t^2} \quad \text{and} \quad (2\pi)^{-\frac{1}{2}k}[-\rho^2 F_D+\chi_D],$$

respectively, so that

$$\int_D \left|\frac{\partial^2 f}{\partial t^2}-\nabla^2 f\right|^2 dx = \frac{1}{(2\pi)^k}\int \left|\frac{\partial^2 F_D}{\partial t^2}+\rho^2 F_D-\chi_D\right| du = 0,$$

and, since $\partial^2 f/\partial t^2$ and $\nabla^2 f$ are continuous, it follows that they are equal everywhere.

4. Even values of k

The method of solution given above shares with every other method of solution of the wave equation, save that of M. Riesz, the peculiarity that the analysis is different for even values of k from the analysis for odd values of k. For general even values of k the reduction of the solution from the form (12) to the form (17) is very difficult, and it is simplest to rely on Hadamard's 'méthode de descente', by which the

solution for $k = 2m$ can be deduced from that for $k = (2m+1)$. In this method, we substitute for the problem of solving

$$\frac{\partial^2 f}{\partial x_1^2}+\dots+\frac{\partial^2 f}{\partial x_k^2} = \frac{\partial^2 f}{\partial t^2},$$

with

$$f_t(x_1,\dots,x_k,0) = \phi(x_1,\dots,x_k), \qquad f(x_1,\dots,x_k,0) = \psi(x_1,\dots,x_k)$$

the problem of solving

$$\frac{\partial^2 F}{\partial x_1^2}+\dots+\frac{\partial^2 F}{\partial x_{k+1}^2} = \frac{\partial^2 F}{\partial t^2},$$

with

$$F(x_1,\dots,x_{k+1},0) = \phi(x_1,\dots,x_k), \qquad F_t(x_1,\dots,x_{k+1},0) = \phi(x_1,\dots,x_k).$$

The solution F is then independent of x_{k+1} and gives the solution of the original problem, which is found to take the same form (17) as above.

These difficulties apply only to the problem of getting the solution into the final form (17); the proof that the value of the solution at a point depends only on the initial values within a sphere of radius t about the point is as before. We may add that for the case $k = 2$ the solution by the direct method presents no difficulty. Consider the first term in (12), where for $k = 2$ we get

$$\begin{aligned} H_\sigma(r) &= 2\pi \int_0^\infty e^{-\delta\rho} \sin \rho t \, J_0(r\rho)\, d\rho \\ &= 2\pi \int_r^\infty \left[\frac{\delta}{\delta^2+(t-v)^2} - \frac{\delta}{\delta^2+(t+v)^2}\right] \frac{dv}{\sqrt{(v^2-r^2)}}, \end{aligned}$$

using Parseval's formula with a formula for the sine transform of J_0. The first term in (12) then gives

$$\begin{aligned} &\frac{1}{2\pi}\int_0^\infty r\phi_x(r)\,dr \int_r^\infty \left[\frac{\delta}{\delta^2+(t-v)^2} - \frac{\delta}{\delta^2+(t+v)^2}\right] \frac{dv}{\sqrt{(v^2-r^2)}} \\ &\qquad = \frac{1}{2\pi}\int_0^\infty \left[\frac{\delta}{\delta^2+(t-v)^2} - \frac{\delta}{\delta^2+(t+v)^2}\right] dv \int_0^r \frac{r\phi_x(r)}{\sqrt{(v^2-r^2)}}\,dr \\ &\qquad \to \int_0^t \frac{r\phi_x(r)}{\sqrt{(t^2-r^2)}}\,dr \quad (\delta \to 0). \end{aligned}$$

In the same way, the second term in (12) gives the other term in (17); here ϕ must be assumed differentiable and ψ twice differentiable.

5. Differentiability and propagation of the solution

The values of f_t and of f for $t = t_0$ constitute data for a new initial-value problem, and, for any $t > t_0$, the solution must be obtainable from that at $t = t_0$ in the same way as it is from that at $t = 0$; this is the essence of Huygens's principle. It is necessary for the validity of this method that f should satisfy the appropriate differentiability conditions at $t = t_0$. Now the Hadamard solution of the wave equation, like that above, requires that the initial values of the function have derivatives of order $[\frac{1}{2}k]+1$; the existence of derivatives of this order at $t = t_0$ requires the existence of derivatives of twice this order at $t = 0$: and thus to maintain Huygens's principle we should appear to require that the initial values be analytic; for otherwise the interposition of a sufficient number of intermediate values of t between 0 and the final value would cause the necessary differentiability assumptions to fail. We shall show that, if the derivatives of a given order of the initial values of f and f_t are integrable in square over any bounded set, then so are the derivatives of the same order at any later value of t. The differentiability conditions, in this stronger form, are *transmitted*. [Cf. **9**, where other transmitted (*fortsetzbare*) conditions are given for $k = 2$, and also **1**, 468–9.]

We have shown that the values of f and of its derivatives inside a fixed sphere D' at time t depend only on the initial values of f and f_t inside a sphere D'', with the same centre as D' and radius greater than that of D' by t. This proof required no assumptions of differentiability. In order to discuss the summability over D' of f and its derivatives at time t, we may therefore alter the initial-value problem by taking, in place of ψ and ϕ, functions equal to them inside D'', vanishing outside a sphere D concentric with D'' but of larger radius, and continuous, with derivatives of all the necessary orders (i.e. up to the order of the highest derivative being considered) existing everywhere and of summable square over D. The function which is the solution of this new initial-value problem has values at time t coinciding with those of f everywhere inside D': and, if we show that this new function has derivatives of the appropriate orders which are of summable square over the whole space, we shall have shown, *a fortiori*, that those of f are summable in square over D'. This amounts to saying that in order to prove our result we need only consider the case where the initial values vanish outside some finite sphere, and hence we shall suppose this to be so and shall therefore extend the appropriate integrals over the whole space in what follows, since the integrands vanish outside a finite region for any value of t.

If the Fourier transform of f is

$$F(u_1,\ldots,u_k) = \frac{1}{(2\pi)^k}\int f(x)e^{iu.x}\,dx,$$

then, on integrating by parts, the Fourier transform of $\partial^r f/\partial x_j^r$ is $(-iu_j)^{-r}F(u)$ if $\partial^r f/\partial x_j^r$ is of summable square over the whole space: then so, by the ordinary L^2 theory, is $F/(-iu_j)^r$. Since the Fourier transform of f is

$$\Phi(u)\frac{\sin\rho t}{\rho}+\Psi(u)\cos\rho t, \tag{18}$$

and that of $\partial^r f/\partial x_j^r$ is $(-iu_j)^{-r}$ times this expression, we see that, if $\dfrac{\partial^{r-1}\phi}{\partial x_j^{r-1}}$ and $\dfrac{\partial^r\psi}{\partial x_j^r}$ are of summable square then so are $(-iu_j)^{-r+1}\Phi$ and $(-iu_j)^r\Psi$, and hence so is $(-iu_j)^{-r}F$, and consequently so is $\partial^r f/\partial x_j^r$. Applying a similar argument to f_t we get the following result:

If all the s-order differential coefficients of $f_t(x,0)$ and all the $(s+1)$-order differential coefficients of $f(x,0)$ are of summable square over any bounded set, then the same is true of the corresponding differential coefficients of $f_t(x,t)$ and $f(x,t)$ for any t.

The existence of a function $f(x,t)$ for every point x can only be inferred if $s \geqslant [\frac{1}{2}k]+1$; and even then the function $f(x,t)$ need not be differentiable.

However, for any $s \geqslant 0$, the function whose Fourier transform is (18) does exist in a mean-square sense, and its derivatives of order s also exist almost everywhere as functions of summable square. The function obtained in this way is a solution of the wave equation, in a generalized sense, which may be defined as follows.

It is clear that, if $f(x,t)$ satisfies the wave equation in the ordinary sense over a bounded interval of values of x and t, then, if $g(x,t)$ is any function twice differentiable with respect to x and t, and vanishing with its derivatives of first order outside and on the boundary of the interval of values of x and t, we shall have

$$\int f\left(\nabla^2 g-\frac{\partial^2 g}{\partial t^2}\right)dxdt = 0, \tag{19}$$

where the integration is taken over the given interval of x and t. Conversely, if a twice differentiable function $f(x,t)$ satisfies (19) for every twice differentiable function $g(x,t)$ of the type considered, then $f(x,t)$ is a solution of the wave equation.

We may call a general function $f(x,t)$ a 'solution of the wave equation' if it satisfies (19) for every finite interval in the (x,t) space, and for every

twice differentiable $g(x,t)$ vanishing with its derivatives of first order on the boundary of that interval. It is then clear that the Fourier transform of (18) is a solution of the wave equation in this generalized sense; and, if its initial values have second-order derivatives, it will be a solution in the ordinary sense. Another way of looking at these generalized solutions is that provided by generalized functions, or 'distributions', of Laurent Schwartz; in the set of generalized functions, every function has derivatives of all orders, and the differential equation can be written down for any function. However, as the definition of the 'distributions' depends essentially on repeated integrations by parts, this point of view does not differ fundamentally from that here adopted.

REFERENCES

1. Courant und Hilbert, *Methoden der Mathematischen Physik* (Springer), vol. ii (1937).
2. N. Zeilon, *Arkiv for Math. Astr. Fys*, 6 (1912), Nr. 38.
3. M. Heins, *J. of Math. and Phys.* (Massachusetts Inst. Tech.), 14 (1935), 137–42.
4. J. L. B. Cooper, *J. of London Math. Soc.* 14 (1939), 124–8.
5. S. Bochner, *Trans. American Math. Soc.* 40 (1936), 175–207.
6. —— *Vorlesungen über Fourierscher Integrale* (Leipzig, 1932).
7. E. C. Titchmarsh, *Theory of Fourier Integrals* (Oxford, 1937).
8. J. Hadamard, *Le Problème de Cauchy* (Paris, 1932).
9. K. Friedrichs and H. Lewy, *Nachr. Ges. Wiss. Göttingen* (1932), 135–43.

ON SOME DIVISOR SUMS ASSOCIATED WITH DIOPHANTINE EQUATIONS

By RICHARD BELLMAN (*Stanford*)

[Received 4 April 1949]

1. Introduction

In this paper I consider the problem of finding an asymptotic expression for each of the sums

$$S_3 = \sum_{n<N} d(n)d_3(n+l), \qquad S_4 = \sum_{n<N} d(n)d_4(n+l), \tag{1}$$

where $d(n)$ is the divisor function, and $d_3(n)$, $d_4(n)$ are defined by the series

$$\zeta(s)^3 = \sum_{n=1}^{\infty} d_3(n)n^{-s}, \qquad \zeta(s)^4 = \sum_{n=1}^{\infty} d_4(n)n^{-s} \quad \{\mathrm{re}(s) > 1\}, \tag{2}$$

or by the arithmetic equivalents

$$d_3(n) = \sum_{x_1x_2x_3=n} 1, \qquad d_4(n) = \sum_{x_1x_2x_3x_4=n} 1. \tag{3}$$

The sums S_3, S_4 arise if we ask the number of solutions in positive rational integers of the respective Diophantine equations

$$\left.\begin{array}{ll} (a) & x_1x_2x_3 - y_1y_2 = l \quad (y_1y_2 \leqslant N), \\ (b) & x_1x_2x_3x_4 - y_1y_2 = l \quad (y_1y_2 \leqslant N). \end{array}\right\} \tag{4}$$

The corresponding sum

$$S_2 = \sum_{n<N} d(n)d(n+l) \tag{5}$$

arising from the Diophantine equation

$$x_1x_2 - y_1y_2 = l \tag{6}$$

was first treated by Ingham [**4**] and subsequently by Estermann [**2**]. The circle method of Hardy and Littlewood was used by Titchmarsh [**6**] to obtain a heuristic formula for S_3.

My principal result is

Theorem 1. *As* $N \to \infty$

$$\left.\begin{array}{l} \sum_{n<N} d(n)d_3(n+l) \sim c_3(l)N\log^3 N \\ \sum_{n<N} d(n)d_4(n+l) \sim c_4(l)N\log^4 N \end{array}\right\}, \tag{7}$$

where $c_3(l)$ and $c_4(l)$ are (complicated) functions of l.

The proof of Theorem 1 is more complicated in detail than in principle, and to illustrate the method I shall therefore restrict myself to

the case $l = 1$, where the computation is at a minimum. After the proof of this result I shall discuss the general case.

2. Reduction of the problem

In this section I show that the estimation of S_3 and S_4 can be reduced to the problem of obtaining asymptotic expressions for

$$T_3 = \sum_{n<N} d_3(kn+1), \qquad T_4 = \sum_{n<N} d_4(kn+1), \tag{1}$$

with sufficiently accurate error terms, which (we note) depend upon N and k. As far as the dependence on N is concerned, we experience no difficulty, and standard methods would suffice to establish more precise results in this respect than I obtain here. It is the necessity for obtaining simultaneous bounds in both N and k that creates the hardships.

LEMMA 1. *If, as* $N \to \infty$,

$$\left.\begin{aligned} T_3 &= N\log^2 Nk\,\psi_1(k) + O(N\log Nk) + O\{a(k)\sqrt{(Nk)}[\log\sqrt{(Nk)}]^{3-\epsilon}\}, \\ T_4 &= N\log^3 Nk\,\psi_2(k) + O(N\log^2 Nk) + O\{a(k)\sqrt{(Nk)}[\log\sqrt{(Nk)}]^{4-\epsilon}\}, \end{aligned}\right\} \tag{2}$$

for some positive constant ϵ, *where the constants implied by the* O'*s are independent of* N *and* k, *and* $a(k)$, $\psi_1(k)$, $\psi_2(k)$ *are functions of* k *with constant mean values, then, as* $N \to \infty$,

$$\left.\begin{aligned} \sum_{n<N} d(n)d_3(n+1) &\sim c_1 N\log^3 N \\ \sum_{n<N} d(n)d_4(n+1) &\sim c_2 N\log^4 N \end{aligned}\right\}. \tag{3}$$

Proof. I use the following important symmetric property of $d(n)$:

$$d(n) = \sum_{k|n} 1 = 2\sum_{\substack{k|n \\ k^2<n}} 1 + m(n), \tag{4}$$

where

$$m(n) = \begin{cases} 1 & \text{if } n = r^2, \\ 0 & \text{otherwise.} \end{cases} \tag{5}$$

Hence,

$$\begin{aligned} S_3 = \sum_{n<N} d(n)d_3(n+1) &= \sum_{n<N} d_3(n+1)\Big\{2\sum_{\substack{k|n \\ k^2<n}} 1 + m(n)\Big\} \\ &= 2\sum_{n<N} d_3(n+1)\sum_{\substack{k|n \\ k^2<n}} 1 + \sum_{n<\sqrt{N}} d_3(n^2+1). \end{aligned} \tag{6}$$

The second sum is trivially $O(N^{\frac{1}{2}+\epsilon})$, since $d(n) = O(n^\epsilon)$ for any $\epsilon > 0$, and we turn our attention to the first sum.

Inverting the orders of summation, the result is

$$2\sum_{k=1}^{\sqrt{N}}\Big\{\sum_{\substack{k^2<n<N \\ n\equiv 0(k)}} d_3(n+1)\Big\} = 2\sum_{k=1}^{\sqrt{N}}\Big\{\sum_{k<n<N/k} d_3(nk+1)\Big\}. \tag{7}$$

Using the results in (2), and substituting in the above, the result is

$$2\sum_{k=1}^{\sqrt{N}}\left[\frac{N}{k}\log^2 N\,\psi_1(k)-k\log^2 k\,\psi_1(k)+O\left(\frac{N}{k}\log N\right)+O\{a(k)\sqrt{N}(\log N)^{3-\epsilon}\}\right]$$
$$= c_1 N\log^3 N+O\{N(\log N)^{3-\epsilon}\}. \quad (8)$$

Similarly, we obtain the result for the second sum in (3).

We may therefore turn our attention to the sums T_3 and T_4, and begin with T_4 since the results obtained there are used in the treatment of T_3.

3. The estimation of $T_4 = \sum_{n<N} d_4(nk+1)$

Let $\chi(n)$ range over the $\phi(k)$ characters to modulus k and consider the sum

$$\sum_{n=0}^{\infty}\frac{d_4(1+nk)}{(1+nk)^s} = \frac{1}{\phi(k)}\sum_{\chi}\left\{\sum_{n=1}^{\infty}\frac{\chi(n)d_4(n)}{n^s}\right\}, \quad (1)$$

the identity holding by virtue of the elementary property of characters,

$$\sum_{\chi}\chi(n) = \begin{cases}\phi(k) & \text{when } n\equiv 1 \pmod{k},\\ 0 & \text{otherwise.}\end{cases} \quad (2)$$

Applying the Perron sum-formula to (1), we obtain

$$\sum_{n<N} d(1+nk) = \frac{1}{2\pi i}\sum_{\chi}\frac{1}{\phi(k)}\int_{(2)}\frac{L_\chi^4(s)(1+Nk)^s}{s}\,ds \quad (3)$$

where the (2) below the integral sign denotes integration along the line $2+it$.

If we shift the line of integration to the line $\frac{1}{2}+it$, we obtain a contribution due to the pole of $L_1(s)$ at $s = 1$, $L_1(s)$ being the series corresponding to the principal character. Since

$$L_1(s) = \prod_{p|k}\left(1-\frac{1}{p^s}\right)\zeta(s), \quad (4)$$

we easily find that this residue is

$$N\log^3 N\,\psi_1(k)+O\{N\log^2 N\,\psi_2(k)\}, \quad (5)$$

where the constant implied by O is independent of k, and where $\psi_1(k)$ and $\psi_2(k)$ are arithmetic functions of k which have constant mean values.

Actually, I shall not shift the entire line of integration to $\frac{1}{2}+it$, but shall follow an argument first used by Hardy and Littlewood in connexion with the Piltz divisor problem [**3**]. As was shown by these authors, it is the average order of $|L_\chi(\frac{1}{2}+it)|^4$ that is important, and it is to the mean value

$$\frac{1}{T}\int_0^T |L_\chi(\tfrac{1}{2}+it)|^4\,dt \quad (6)$$

that we turn our attention.

4. The mean value of $|L_\chi(\frac{1}{2}+it)|^4$

The problem of finding the mean value of $|L_\chi(\frac{1}{2}+it)|^4$ is, in its native environment, not simple, and in this case, it is further complicated by the fact that we require an estimate depending upon both T and k.

I shall follow a method due to Titchmarsh [**5**] up to a point, and then at a certain juncture, introduce a procedure I have used [**1**] in the evaluation of the mean value of $|\zeta(\frac{1}{2}+it)|^4$.

Consider the function

$$F(x) = \sum_{n=1}^{\infty} d(n)\chi(n)e^{-nx} \quad \{\mathrm{re}(x) > 0\}. \tag{1}$$

We have
$$\int_0^\infty F(x)x^{s-1}\,dx = \Gamma(s)L_\chi^2(s) \quad \{\mathrm{re}(s) > 1\}. \tag{2}$$

Using the Mellin inversion formula we have

$$F(x) = \frac{1}{2\pi i}\int_{(2)} \Gamma(s)L_\chi^2(s)x^{-s}\,ds. \tag{3}$$

Shifting the line of integration to $\frac{1}{2}+it$ we obtain a residue $R(x)$ due to a pole at $s = 1$ if we are considering the series corresponding to the principal character, and no residue otherwise. This residue has the value

$$\{\log x\,\psi_3(k)+\psi_4(k)\}/x, \tag{4}$$

where $\psi_3(k)$, $\psi_4(k)$ are $O(\log^2 k)$ as $k \to \infty$.

Now set $x = \exp\{u+i(\frac{1}{2}\pi-\delta)\}$ $(0 < \delta < \frac{1}{2}\pi;\ s = \frac{1}{2}+it)$. Thus,

$$e^{\frac{1}{2}u}[F(e^{u+i(\frac{1}{2}\pi-\delta)})-R(e^{u+i(\frac{1}{2}\pi-\delta)})] = \frac{1}{2\pi}\int_{-\infty}^{\infty} \Gamma(\tfrac{1}{2}+it)L_\chi^2(\tfrac{1}{2}+it)e^{-iut-\frac{1}{2}i(\frac{1}{2}\pi-\delta)}\,dt. \tag{5}$$

Applying the Parseval–Plancherel theorem, we obtain

$$L(\delta) = \int_{-\infty}^{\infty} e^u|F(e^{u+i(\frac{1}{2}\pi-\delta)})-R(e^{u+i(\frac{1}{2}\pi-\delta)})|^2\,du$$
$$= \int_{-\infty}^{\infty} |\Gamma(\tfrac{1}{2}+it)|^2e^{(\pi-2\delta)t}|L_\chi(\tfrac{1}{2}+it)|^4\,dt = R(\delta). \tag{6}$$

It follows from the asymptotic expression for $|\Gamma(\frac{1}{2}+it)|^2$, that we have

$$|\Gamma(\tfrac{1}{2}+it)|^2e^{\pi t} \geqslant c_1 \quad (0 \leqslant t < \infty) \tag{7}$$

for some constant c_1. Hence

$$R(\delta) > \int_0^\infty |\Gamma(\tfrac{1}{2}+it)|^2e^{\pi t}e^{-2\delta t}|L_\chi(\tfrac{1}{2}+it)|^4\,dt$$
$$\geqslant c_1\int_0^\infty e^{-2\delta t}|L_\chi(\tfrac{1}{2}+it)|^4\,dt. \tag{8}$$

Once we have obtained an upper bound for the integral, a very simple Tauberian argument, which I present below, yields a bound for

$$\frac{1}{T}\int_0^T |L_\chi(\tfrac{1}{2}+it)|^4\,dt. \tag{9}$$

For this purpose, we must estimate $L(\delta)$. Thus far I have followed Titchmarsh. I now pursue a different path and introduce a functional equation for $F(x)$, an equation which presents an independent interest. A functional equation of this type, for the function

$$D(x) = \sum_{n=1}^{\infty} d(n)\,e^{-nx},$$

was first discovered by Wigert [**7**]. I follow a method due to Landau to derive the functional equation we require. As indicated by Landau in the case of $D(x)$, this functional equation is in some sense a paraphrase of the functional equation for the corresponding zeta function.

5. The approximate functional equation for $F(x)$

Let us first treat the case where χ is a proper character.

We have
$$F(x)-R(x) = \frac{1}{2\pi i}\int_{(\frac{1}{2})} \Gamma(s)L_\chi^2(s)x^{-s}\,ds. \tag{1}$$

Replace s by $1-s$, i.e. t by $-t$, and obtain

$$\begin{aligned} F(x)-R(x) &= \frac{1}{2\pi i}\int_{(\frac{1}{2})} \Gamma(1-s)L_\chi^2(1-s)x^{-(1-s)}\,ds \\ &= \frac{1}{2\pi i x}\int_{(\frac{1}{2})} \Gamma(1-s)\left[\frac{1}{a}\,\frac{\pi^{\frac{1}{2}-\frac{1}{2}s}}{k^{\frac{1}{2}-\frac{1}{2}s}}\,\frac{\Gamma(\frac{1}{2}s)}{\Gamma(\frac{1}{2}-\frac{1}{2}s)}\,L_{\bar\chi}(s)\right]^2 x^s\,ds, \end{aligned} \tag{2}$$

where we have used the functional equation, valid if χ is a proper character,

$$\left(\frac{\pi}{k}\right)^{-\frac{1}{2}s}\Gamma(\tfrac{1}{2}s)L_\chi(s) = a\left(\frac{\pi}{k}\right)^{-(\frac{1}{2}-\frac{1}{2}s)}\Gamma(\tfrac{1}{2}-\tfrac{1}{2}s)L_\chi(1-s), \tag{3}$$

where a is a constant with modulus 1, i.e. $|a| = 1$. Using the duplication formula for $\Gamma(s)$ and performing some elementary simplifications, we have the end result

$$F(x)-R(x) = \frac{1}{2\pi i a^2 kx}\int_{(\frac{1}{2})} \cot\tfrac{1}{2}\pi s\,\Gamma(s)L_{\bar\chi}^2(s)\left(\frac{4\pi}{kx}\right)^{-s} ds. \tag{4}$$

Let us note that, *on any line* $b+it$, where $\sin\frac{1}{2}\pi b \neq 0$,

$$\cot\tfrac{1}{2}\pi s = \begin{cases} i+O(e^{\pi t}) & (t\to-\infty), \\ O(1) & (t\to+\infty). \end{cases} \tag{5}$$

Keeping this in mind, write (4) as

$$\frac{1}{2\pi a^2kx}\int_{(\frac{1}{2})}\Gamma(s)L^2_{\bar{\chi}}(s)\left(\frac{4\pi}{kx}\right)^{-s}ds+\frac{1}{2\pi i a^2kx}\int_{(\frac{1}{2})}(\cot\tfrac{1}{2}\pi s-i)\Gamma(s)L^2_{\bar{\chi}}(s)\left(\frac{4\pi}{kx}\right)^{-s}ds. \tag{6}$$

The first integral is
$$\frac{1}{a^2kx}\left\{\bar{F}\left(\frac{4\pi}{kx}\right)-\bar{R}\left(\frac{4\pi}{kx}\right)\right\} \tag{7}$$

where by $\bar{F}(x)$ we mean the function obtained from $F(x)$ when χ is replaced by its conjugate.

Let us now turn to the second integral. If χ is not the principal character, $L^2_{\bar{\chi}}(s)$ has no singularity at $s = 1$, and we may shift the line of integration to $1+b$, where b is some fixed constant $(0 < b < 1)$, without contributing any residue term. On the line $1+b$, $|L_{\bar{\chi}}(s)|^2$ is uniformly bounded by a constant depending only upon b. If we take x in the form

$$x = |x|e^{i(\frac{1}{2}\pi-\delta)}, \tag{8}$$

where $0 < \delta \leqslant \frac{1}{2}\pi$, we have

$$\left|\left(\frac{4\pi}{kx}\right)^{-s}\right| = |x|^{1+b}k^{1+b}e^{-(\frac{1}{2}\pi-\delta)t}. \tag{9}$$

Now consider the second integral in (7) over the range $(b, b+i\infty)$. Since $|\cot\frac{1}{2}\pi s-i| = 0$ on the half-line of integration, the integral is $O(k^b|x|^b)$ as $|x|\to 0$, $k\to\infty$ since

$$|\Gamma(b+it)| = O(|t|^{b-\frac{1}{2}}e^{-\frac{1}{2}\pi|t|}) \quad \text{as} \quad |t|\to\infty,$$

the constant implied by O being independent of δ.

Over the range $(b, b-i\infty)$ we use $|\cot\frac{1}{2}\pi s-i| = O(e^{\pi t})$ $(t\to-\infty)$, and again we obtain $O(k^b|x|^b)$ as a bound.

If χ is the principal character, we must consider the residue at $s = 1$, which is

$$O\{(\log|x|)\log^2 k\}, \tag{10}$$

as $k\to\infty$, $|x|\to 0$.

Putting these results together, we can state

LEMMA 2. (*Approximate functional equation for* $\sum_{n=1}^{\infty} d(n)\chi(n)e^{-nx}$.) *For* $0 \leqslant \arg x < \frac{1}{2}\pi$, *we have*

$$F(x)-R(x) = \frac{1}{a^2kx}\left\{\bar{F}\left(\frac{4\pi}{kx}\right)-\bar{R}\left(\frac{4\pi}{kx}\right)\right\}+O(k^b|x|^b)+O(\log^2 k\log|x|), \tag{11}$$

where

$$F(x) = \sum_{n=1}^{\infty} d(n)\chi(n)e^{-nx},$$

$$\bar{F}(x) = \sum_{n=1}^{\infty} d(n)\bar{\chi}(n)e^{-nx},$$

$$R(x) = \begin{cases} O\left(\log^2 k \dfrac{\log|x|}{|x|}\right) & \text{as } |x| \to 0, \infty,\ k \to \infty, \\ \quad \chi \text{ a principal character}, \\ 0 & \text{otherwise}, \end{cases} \tag{12}$$

$$\bar{R}(x) = \begin{cases} O\left(\log^2 k \dfrac{\log|x|}{|x|}\right) & \text{as } |x| \to 0, \infty,\ k \to \infty, \\ \quad \bar{\chi} \text{ a principal character}, \\ 0 & \text{otherwise}, \end{cases}$$

and the constants implied by O are independent of k as $k \to \infty$, and of x as $|x| \to 0$, depending only upon b, and are uniform for $\arg x$ in the half-open interval $0 \leqslant \arg x < \frac{1}{2}\pi$.

6. The mean value of $|L_\chi(\frac{1}{2}+it)|^4$ (concluded)

I now use the above result to bound $L(\delta)$, the left-hand side of equation (6) in § 4,

$$L(\delta) = \int_0^\infty |F(xe^{i(\frac{1}{2}\pi-\delta)}) - R(xe^{i(\frac{1}{2}\pi-\delta)})|^2\,dx = \int_0^1 + \int_1^\infty. \tag{1}$$

Let us consider the integral over $(1,\infty)$ first. Since

$$\left(\int_1^\infty |F|^2\,dx\right)^{\frac{1}{2}} - \left(\int_1^\infty |R|^2\,dx\right)^{\frac{1}{2}} \leqslant \left(\int_1^\infty |F-R|^2\,dx\right)^{\frac{1}{2}}$$
$$\leqslant \left(\int_1^\infty |F|^2\,dx\right)^{\frac{1}{2}} + \left(\int_1^\infty |R|^2\,dx\right)^{\frac{1}{2}}, \tag{2}$$

and $\int_1^\infty |R|^2\,dx = O(\log^4 k)$, the principal contribution to the integral is

$$\int_1^\infty |F|^2\,dx = \int_1^\infty \left(\sum_{n=1}^\infty d(n)\chi(n)e^{-nxe^{i(\frac{1}{2}\pi-\delta)}}\right)\left(\sum_{n=1}^\infty d(n)\bar{\chi}(n)e^{-nxe^{-i(\frac{1}{2}\pi-\delta)}}\right)dx. \tag{3}$$

Because of the absolute and uniform convergence of the series in the interval $(1,\infty)$, we may evaluate the product using Cauchy multiplication, and integrate term by term. The resultant double sum has been estimated by Titchmarsh [**5**], who showed that the main contribution is from the cross-product terms

$$\sum_{n=1}^\infty n^{-1}|\chi(n)|^2 d^2(n) e^{-2nx\sin\delta} = O\left\{\frac{1}{\delta}\left(\log\frac{1}{\delta}\right)^3\right\}. \tag{4}$$

Now consider the integral over $(0, 1)$. Using the approximate functional equation derived above, the integral becomes, applying Minkowski's inequality again,

$$\left(\int_0^1\right)^{\frac{1}{2}} = \left\{\int_0^1 \frac{1}{k^2x^2}\left|\bar{F}\left(\frac{4\pi}{kx}e^{-i(\frac{1}{2}\pi-\delta)}\right)-\bar{R}\left(\frac{4\pi}{kx}e^{-i(\frac{1}{2}\pi-\delta)}\right)\right|^2 dx\right\}^{\frac{1}{2}}+O(k^b). \quad (5)$$

Set $kx = y$ and then $y = 1/x$, obtaining

$$\frac{1}{k}\int_{1/k}^{\infty} |F(4\pi xe^{-i(\frac{1}{2}\pi-\delta)})-R(4\pi xe^{-i(\frac{1}{2}\pi-\delta)})|^2\, dx. \quad (6)$$

Going through the previous argument, applying Titchmarsh's procedure once more, we see that this integral is bounded by $O\{(\log k/\delta)^2/\delta\}$. The factor $1/k$ in front of (6) is of crucial importance, since without it we could obtain only $O\{k(\log k/\delta)^2/\delta\}$.

Let us mention that for fixed k the above methods readily yield the asymptotic formula

$$\int_1^T |L_\chi(\tfrac{1}{2}+it)|^4\, dt \sim c_1(k)T(\log T)^3 \quad (7)$$

as $T \to \infty$.

Putting the above results together, we obtain, finally,

$$L(\delta) = O\{(\log k/\delta)^3/\delta\}, \quad (8)$$

where the constant implied by O is independent of k, as $k \to \infty$, and of δ as $\delta \to 0$. Consequently

$$e^{-1}\int_0^{1/\delta} |L_\chi(\tfrac{1}{2}+it)|^4\, dt \leqslant \int_0^{1/\delta} e^{-\delta t}|L_\chi(\tfrac{1}{2}+it)|^4\, dt = O\left\{\frac{1}{\delta}\left(\log\frac{k}{\delta}\right)^3\right\}, \quad (9)$$

i.e.

$$\int_0^T |L_\chi(\tfrac{1}{2}+it)|^4\, dt = O\{T(\log kT)^3\}. \quad (10)$$

From (10), using Holder's inequality, we obtain

$$\int_0^T |L_\chi(\tfrac{1}{2}+it)|^3\, dt = O\{T(\log kT)^{\frac{9}{4}}\}. \quad (11)$$

This estimate could be improved, but the fact that $\frac{9}{4} < 3$ is all that we require for the proof.

7. Estimation of the remainder term

We now turn to (3) of § 3, and the estimation of the remainder term. In place of the contour $2+it$ I use the contour consisting of the five

line segments $(2+iT,\infty)$, $(2-iT,-\infty)$, $(\frac{1}{2}+iT,2+iT)$, $(\frac{1}{2}-iT,2-iT)$, $(\frac{1}{2}-iT,\frac{1}{2}+iT)$, where T will be chosen in a convenient manner subsequently. We then have

$$\int_{(2)} = \int_{2+iT}^{2+i\infty} + \int_{\frac{1}{2}+iT}^{2+iT} + \int_{\frac{1}{2}-iT}^{\frac{1}{2}+iT} + \int_{2-iT}^{\frac{1}{2}-iT} + \int_{2-i\infty}^{2-iT} = I_1+I_2+I_3+I_4+I_5. \tag{1}$$

Let us first treat the case where χ is a proper, but not principal character. I shall then note the changes necessary when χ is not a proper character. Consider I_3 first. We have

$$|I_3| \leqslant \left(2\int_{-T}^{T} \frac{|L_\chi(\frac{1}{2}+it)|^4}{1+|t|}(Nk)^{\frac{1}{2}}\,dt\right) = O\{(Nk)^{\frac{1}{2}}(\log kT)^3\}, \tag{2}$$

using (8) of the previous section. Consider I_2. First of all, we can easily show that (8) implies

$$\int_{-T}^{T} |L_\chi(a+it)|^4\,dt = O\{T(\log kT)^3\} \quad (\tfrac{1}{2} \leqslant a < \infty), \tag{3}$$

with the constant implied by O uniform in a. Hence

$$\int_{\frac{1}{2}}^{2}\left(\int_{0}^{T} |L_\chi(a+it)|^4\,dt\right)da = O\{T(\log kT)^3\}. \tag{4}$$

Therefore there is at least one T_1 between T and $2T$ such that

$$\int_{\frac{1}{2}}^{2} |L_\chi(a+iT_1)|^4\,da = O\{(\log kT)^3\}, \tag{5}$$

and this is the variable that we denote by T itself. It follows that

$$|I_2| = O\{(Nk)^2(\log kT)^3/T\}. \tag{5 a}$$

Finally consider I_1. It is easy to see that the mean-value theorem, applied to the real and imaginary parts separately, and the absolute convergence of the series for $L_\chi(2+it)$, combine to yield the estimate

$$|I_1| = O\left\{\frac{(Nk)^2}{T}\right\}. \tag{6}$$

The remaining integrals I_4, I_5 correspond to I_2 and I_1 respectively, and, combining the above estimates, the result is

$$\left|\int_{(2)} L_\chi^4(s)(Nk)^s\frac{ds}{s}\right| = O\{(Nk)^{\frac{1}{2}}(\log kT)^3\}+O\left\{(\log kT)^3\frac{(Nk)^2}{T}\right\}. \tag{7}$$

Now choose $T = (Nk)^2$. Then

$$\left|\int\limits_{(2)} L_{\chi}^{4}(s)(Nk)^{s}\frac{ds}{s}\right| = O\{(Nk)^{\frac{1}{2}}(\log Nk)^{3}\}. \tag{8}$$

If χ is the principal character, the same error term is obtained, together with the principal term due to the residue at $s = 1$.

Precisely the same methods yield the corresponding result

$$\left|\int\limits_{(2)} L_{\chi}^{3}(s)(Nk)^{s}\frac{ds}{s}\right| = O\{(Nk)^{\frac{1}{2}}(\log Nk)^{\frac{3}{2}}\}. \tag{9}$$

Now let us discuss the modifications necessary when χ is not a proper character. In this case, we have

$$L_{\chi}(s) = \prod_{\substack{p|k \\ p\nmid K}} (1-\chi(p)/p^{s}) \sum_{n=1}^{\infty} X(n)/n^{s}, \tag{10}$$

where $X(n)$ is a proper character to modulus K, K being the integer with the property that $\chi(n) = \chi(n')$, $n \equiv n' (\bmod K)$.

The first factor on the right in (10) is bounded on $s = \frac{1}{2}+it$ by

$$\psi(k) = \prod_{p|k} (1+1/p^{\frac{1}{2}}), \tag{11}$$

and it is easy to verify that

$$\sum_{k<N} \psi(k) \sim c_1 N. \tag{12}$$

Collecting the previous results, we can state

THEOREM 2.

$$\left.\begin{aligned} \sum_{n<N} d_3(kn+1) &\sim \psi_1(k)\, N \log^2 Nk + O\{a(k)\sqrt{(Nk)}(\log Nk)^{3-\epsilon}\} \\ \sum_{n<N} d_4(kn+1) &\sim \psi_2(k)\, N \log^3 Nk + O\{a(k)\sqrt{(Nk)}(\log Nk)^{4-\epsilon}\} \end{aligned}\right\}, \tag{13}$$

for some positive constant ϵ, *where the constants implied by the O's are independent of* N *and* k, *and* $a(k)$, $\psi_1(k)$, $\psi_2(k)$, *are functions of* k *with constant mean values.*

8. The general sums $\sum_{n<N} d_3(kn+l)$, $\sum_{n<N} d_4(kn+l)$

In this section I briefly turn to the treatment of the general sums

$$\sum_{n<N} d(n)d_3(n+l), \qquad \sum_{n<N} d(n)d_4(n+l). \tag{1}$$

The situation here is, as before, dependent upon the simpler sums

$$\sum_{n<N} d_3(kn+l), \qquad \sum_{n<N} d_4(kn+l). \tag{2}$$

Let us consider the sum

$$\sum_{n<N} d_3(kn+l) = \sum_{n<N} d_3\left\{(k,l)\left(\frac{nk}{(k,l)}+\frac{l}{(k,l)}\right)\right\}. \tag{3}$$

Using the characters to modulus $k/(k,l)$, we obtain, as above, the identity

$$\sum_{n=0}^{\infty} d_3(kn+l)/(kn+l)^s = \frac{1}{\phi\{k/(k,l)\}} \sum_{\chi} \frac{1}{\chi\{l/(k,l)\}} \left\{\sum_{n=1}^{\infty} \frac{\chi(n)d_3\{n(k,l)\}}{n^s}\right\}. \tag{4}$$

We see then that the problem reduces to evaluating the Dirichlet series

$$\sum_{n=1}^{\infty} \chi(n)d_3(an)n^{-s}, \tag{5}$$

in terms of powers of $L_\chi(s)$. This may be done, either by straightforward drudgery, using

$$d_3(an) = \sum_{k|an} d(k), \tag{6}$$

writing
$$\sum_{n=1}^{\infty} \chi(n)d_3(an)n^{-s} = \sum_{l=1}^{\infty} d(l)\left\{\sum_{an\equiv 0(l)} \chi(n)/n^s\right\}, \tag{7}$$

and continuing in this vein, or perhaps a bit more simply by using a generalization of a Ramanujan–Bachmann formula expressing $d(uv)$ in terms of $d(u/k)$, $d(v/k)$, k running through the divisors of u and v.

In any case, it is soon seen that the amount of effort required to evaluate the various constants occurring, which depend on l, is excessive and prohibitive, considering the result derived.

REFERENCES

1. R. Bellman, 'The mean-value of $\zeta(s)^4$ on the critical line', *Duke Math. J.* (in the press).
2. T. Estermann, 'On the representation of a number as a sum of three products', *Proc. London Math. Soc.* 29 (1929), 453–78.
3. G. H. Hardy and J. E. Littlewood, 'The approximate functional equation in the theory of the zeta function with applications to the divisor problems of Dirichlet and Piltz', ibid. (2) 21 (1922), 39–74.
4. A. E. Ingham, 'Some asymptotic formulae in the theory of numbers', *J. of London Math. Soc.* 2 (1927), 207–8.
5. E. C. Titchmarsh, 'The mean-value of the zeta function on the critical line', *Proc. London Math. Soc.* (2) 27 (1928), 137–50.
6. —— 'Some problems in the analytic theory of numbers', *Quart. J. of Math.* (Oxford) 13 (1942), 129–52.
7. S. Wigert, 'Sur la série de Lambert et son application à la théorie des nombres', *Acta Math.* 41 (1916), 197–218.

CANONICAL FORMS (II): PARALLEL PARTIALLY NULL PLANES

By A. G. WALKER (*Sheffield*)

[Received 7 April 1949]

IN a recent paper† I gave a canonical form for the general Riemannian n-space that admits a parallel null r-plane. The present paper is concerned with the more general problem when the parallel plane is partially but not necessarily wholly null, and I give a canonical form for the general V_n that admits a parallel plane of dimensionality $r+s$ and nullity r. By 'nullity r' is meant that the plane has a null part of dimensionality r, this null part being the intersection of the $(r+s)$-plane and its conjugate (orthogonal) $(n-r-s)$-plane. Since the conjugate plane contains the null part, $n-r-s \geqslant r$, i.e.

$$2r+s \leqslant n.$$

The present canonical form reduces to that of C.F.(I) when $s = 0$ and again when $2r+s = n$; in the former case the given plane is wholly null, and in the latter case the conjugate plane is wholly null.

As before, we suppose that the space is either of class C^∞ or of class C^ω, and we consider only the local problem.

The $n \times n$ matrices occurring in this paper will usually be expressed in terms of sub-matrices according to the partition of n

$$n = (r)+(s)+(n-2r-s)+(r) \tag{1}$$

for both rows and columns. Suffix sets corresponding to this partition will be denoted as follows:

$$L = (1,2,\dots,r), \qquad M = (r+1,\dots,r+s),$$

$$N = (r+s+1,\dots,n-r), \qquad L' = (n-r+1,\dots,n).$$

Latin suffixes will take values 1, 2,..., n throughout, and Greek suffixes will take values as indicated.

The general theorem is as follows:

† See above, 70–9; this paper will be referred to as C.F.(I). The terminology and notation of the present paper are those of C.F.(I) and of a previous paper *Quart. J. of Math.* 20 (1949), 135–45.

Theorem. *A canonical form for the general V_n that admits a parallel $(r+s)$-plane of nullity r is given by the fundamental tensor*

$$(g_{ij}) = \begin{pmatrix} O & O & O & I \\ O & A & O & F \\ O & O & B & G \\ I & F' & G' & C \end{pmatrix}, \tag{2}$$

where the non-zero sub-matrices satisfy the following conditions:

(i) *orders are given by the partition* (1), *I is the unit $r\times r$ matrix, A and B are symmetric and non-singular, C is symmetric, and F', G' are the transposes of F, G respectively;*

(ii) *A and F (and so also F') are independent of the coordinates x^α $(\alpha \in L+N)$; and B and G (and so also G') are independent of x^β $(\beta \in L+M)$.*

A basis for the parallel $(r+s)$-plane is provided by the vectors δ^i_1, δ^i_2,..., δ^i_{r+s} at each point of V_n. The first r of these vectors also provide a basis for its null part, the parallel null r-plane.

Let V_n be a Riemannian n-space that admits a parallel $(r+s)$-plane p of nullity r; I shall prove that a coordinate system exists for which the fundamental tensor g_{ij} of V_n takes the form (2) with the conditions of the theorem satisfied. We write p' for the parallel $(n-r-s)$-plane conjugate to p, and p^* for the parallel r-plane which is the null part of p. The parallel plane conjugate to p^* is the $(n-r)$-plane $p+p'$. Since p^* is the intersection of p and p', we may choose basis vectors $\lambda^i_{(\alpha)}$ so that they form a basis of p^* when $\alpha \in L$, of p when $\alpha \in L+M$, of p' when $\alpha \in L+N$, and so of $p+p'$ when $\alpha \in L+M+N$.

Writing X_α for the linear operator $\lambda^i_{(\alpha)}\partial_i$, then [C.F.(I) § 2] to each parallel plane there corresponds a complete system of partial differential equations

$$X_\alpha f = 0. \tag{3}$$

The independent solutions of these equations will be written $f^{(\rho)}$ where

$\rho \in L'$	when $\alpha \in L+M+N$	(corresponding to $p+p'$);	(4 a)
$\rho \in N+L'$	,, $\alpha \in L+M$	(,, ,, p);	(4 b)
$\rho \in M+L'$	,, $\alpha \in L+N$	(,, ,, p');	(4 c)
$\rho \in M+N+L'$	,, $\alpha \in L$	(,, ,, p^*).	(4 d)

The functions $f^{(\rho)}$ for $\rho \in L'$ are independent and satisfy

$$\lambda^i_{(\alpha)} f^{(\rho)}{}_{,i} = 0 \quad (\alpha \in L+M+N).$$

The r covariant vectors $f^{(\rho)}{}_{,i}$ $(\rho \in L')$ are therefore independent and orthogonal to the vectors of $p+p'$, i.e. they form a basis for p^*. We may therefore take

$$\lambda^i_{(\alpha)} = g^{ij} f_{(\alpha),j}, \qquad f_{(\alpha)} = f^{(\alpha+n-r)} \quad (\alpha \in L). \tag{5}$$

It can be verified as in C.F.(I) that with this choice of basis for p^* the corresponding operators satisfy

$$X_\alpha X_\beta - X_\beta X_\alpha = 0 \quad (\alpha, \beta \in L).$$

Hence, by the lemma of C.F.(I) § 3, the system of equations for f

$$X_\alpha f = c_\alpha \quad (\alpha \in L),$$

where c_α are any constants, admit a solution. In particular there are functions $f^{(\beta)}$ $(\beta \in L)$ which satisfy

$$X_\alpha f^{(\beta)} = \delta^\beta_\alpha \quad (\alpha, \beta \in L). \tag{6}$$

Combining these equations with those given by (3) and (4 d), we have

$$X_\alpha f^{(i)} = \delta^i_\alpha \quad (\alpha \in L). \tag{7}$$

We have now introduced n functions $f^{(i)}$; $n-r$ of them by (3) and (4), and the remaining r by (6). It can easily be verified as in C.F.(I) that these functions are independent, and we may therefore transform to new coordinates x'^i where

$$x'^i = f^{(i)}.$$

In the new coordinate-system, omitting the primes, $f^{(i)} = x^i$, so that $X_\alpha f^{(i)} = \lambda^i_{(\alpha)}$, and (7) becomes

$$\lambda^i_{(\alpha)} = \delta^i_\alpha \quad (\alpha \in L). \tag{8}$$

The remaining equations given by (3) and (4) are now

$$X_\mu x^\rho = \lambda^\rho_{(\mu)} = 0 \quad (\mu \in M;\ \rho \in N+L'); \tag{9 a}$$

$$X_\nu x^\sigma = \lambda^\sigma_{(\nu)} = 0 \quad (\nu \in N;\ \sigma \in M+L'). \tag{9 b}$$

We see from (8) that the r vectors δ^i_α $(\alpha \in L)$ form a basis for p^*; from (8) and (9 a) that the $r+s$ vectors δ^i_μ $(\mu \in L+M)$ are contained in the $(r+s)$-plane p; and from (8) and (9 b) that the $n-r-s$ vectors δ^i_ν $(\nu \in L+N)$ are contained in the $(n-r-s)$-plane p'. Since any set of t independent vectors contained in a t-plane may be taken to form a basis, we may assume

$$\lambda^i_{(\alpha)} = \delta^i_\alpha \quad (\alpha \in L+M+N). \tag{10}$$

Using the fact that p and p' are orthogonal we have from (10)

$$g_{ij}\lambda^i_{(\mu)}\lambda^j_{(\nu)} = g_{\mu\nu} = 0 \quad (\mu \in L+M;\ \nu \in L+N). \tag{11}$$

From (5), $\quad g_{ij}\lambda^j_{(\alpha)} = f^{(\alpha')}{}_{,i} \quad (\alpha \in L;\ \alpha' = \alpha+n-r),$

and, since $f^{(\alpha')} = x^{\alpha'}$ and $\lambda^j_{(\alpha)} = \delta^j_\alpha$, these relations become

$$g_{i\alpha} = \delta^{\alpha+n-r}_i \quad (\alpha \in L). \tag{12}$$

From (11) and (12) it follows that the matrix (g_{ij}) takes the form

$$\begin{pmatrix} O & O & O & I \\ O & A & O & F_1 \\ O & O & B & G_1 \\ I & F_1' & G_1' & C_1 \end{pmatrix}, \tag{13}$$

where I, A, B, C, F_1, G_1, F_1', G_1' are sub-matrices satisfying the conditions (i) of the theorem, A and B being non-singular because

$$||g_{ij}|| = (-1)^r ||A||.||B|| \neq 0.$$

We now use the fact that p is parallel, the conditions for which are

$$\lambda^i_{(\alpha),k} = A^\beta_{\alpha k} \lambda^i_{(\beta)} \quad (\alpha, \beta \in L+M)$$

for some $A^\beta_{\alpha k}$. With $\lambda^i_{(\alpha)} = \delta^i_\alpha$ these conditions become

$$\begin{Bmatrix} i \\ \alpha k \end{Bmatrix} = A^\beta_{\alpha k} \delta^i_\beta \quad (\alpha, \beta \in L+M)$$

where $\begin{Bmatrix} i \\ jk \end{Bmatrix}$ are the Christoffel symbols of the second kind. The equations with $i \in L+M$ merely determine $A^\beta_{\alpha k}$; the conditions we are seeking will therefore be given by writing $i = \rho \in N+L'$. These conditions are

$$\begin{Bmatrix} \rho \\ \alpha k \end{Bmatrix} = 0 \quad (\alpha \in L+M;\ \rho \in N+L'). \tag{14}$$

Some of the Christoffel symbols of the first kind are

$$[\mu\alpha k] = g_{\mu\beta}\begin{Bmatrix} \beta \\ \alpha k \end{Bmatrix} + g_{\mu\rho}\begin{Bmatrix} \rho \\ \alpha k \end{Bmatrix} \quad \begin{pmatrix} \alpha, \beta \in L+M, \\ \mu \in L+N;\ \rho \in N+L' \end{pmatrix},$$

and, since $g_{\mu\beta} = 0$ from (13) and $\begin{Bmatrix} \rho \\ \alpha k \end{Bmatrix} = 0$ from (14), we have

$$[\mu\alpha k] = 0 \quad (\alpha \in L+M;\ \mu \in L+N). \tag{15}$$

Conversely it can easily be verified that (14) follows from (15), so that (14) and (15) are equivalent sets of conditions when the fundamental tensor is of the form (13).

We thus have

$$[\mu\alpha k] = \tfrac{1}{2}(\partial_k g_{\alpha\mu} + \partial_\alpha g_{\mu k} - \partial_\mu g_{\alpha k}) = 0 \quad \begin{pmatrix} \alpha \in L+M \\ \mu \in L+N \end{pmatrix},$$

and, since $g_{\alpha\mu} = 0$ from (13), these conditions become

$$\partial_\alpha g_{\mu k} = \partial_\mu g_{\alpha k} \quad (\alpha \in L+M;\ \mu \in L+N). \tag{16}$$

Taking $\alpha \in L$, then, from (13), $g_{\alpha k} = 0$ or 1, so that $\partial_\mu g_{\alpha k} = 0$. We therefore have, from (16) with $\alpha \in L$,

$$\partial_\alpha g_{\mu k} = 0 \quad (\alpha \in L;\ \mu \in L+N),$$

showing that B, G_1, G_1' in (13) are independent of x^α $(\alpha \in L)$. Similarly, taking $\mu \in L$ in (16), we find that A, F_1, F_1' in (13) are also independent of x^α $(\alpha \in L)$.

Taking $\alpha \in M$, $\mu \in N$ and $k = \beta \in M$ in (16), then $g_{\mu\beta} = 0$ from (13) and we have

$$\partial_\mu g_{\alpha\beta} = 0 \quad (\alpha, \beta \in M;\ \mu \in N).$$

Hence A in (13) is independent of x^μ $(\mu \in N)$. Similarly, taking $\alpha \in M$, $\mu \in N$, and $k = \nu \in N$ in (16), we find that B in (13) is independent of x^α $(\alpha \in M)$.

We have now established conditions (ii) of the theorem except for showing that the coordinate-system may be chosen so that F is independent of x^μ $(\mu \in N)$ and G is independent of x^α $(\alpha \in M)$.

From conditions (16) with $\alpha \in M$, $\mu \in N$, $k = \tau \in L'$, we deduce that $g_{\rho\tau}$ $(\rho \in M+N;\ \tau \in L')$ is expressible in the form

$$g_{\rho\tau} = \partial_\rho \theta_\tau + h_{\rho\tau} \quad (\rho \in M+N;\ \tau \in L') \tag{17}$$

for some functions θ_τ independent of x^α $(\alpha \in L)$, and some functions $h_{\rho\tau}$, where $h_{\alpha\tau}$ $(\alpha \in M;\ \tau \in L')$ are independent of x^β $(\beta \in L+N)$ and $h_{\mu\tau}$ $(\mu \in N;\ \tau \in L')$ are independent of x^γ $(\gamma \in L+M)$. I shall write F for the $s \times r$ matrix $(h_{\alpha\tau})$ where $\alpha \in M$, $\tau \in L'$; G for the $(n-2r-s)\times r$ matrix $(h_{\mu\tau})$ where $\mu \in N$, $\tau \in L'$; and K for the $r \times r$ matrix $(\partial_\sigma \theta_\tau)$ where $\sigma, \tau \in L'$.

Consider the transformation from x^i to new coordinates X^i where

$$X^\alpha = x^\alpha - \theta_{\alpha+n-r}, \qquad X^\nu = x^\nu \quad (\alpha \in L;\ \nu \in M+N+L'). \tag{18}$$

This is permissible in that it does not affect any of the previous results; the functions $f^{(\alpha)}$ $(\alpha \in L)$ giving rise to x^α were not defined uniquely as solutions of (6), and to each may be added any solution of $X_\alpha f = 0$ $(\alpha \in L)$, i.e. any function of x^ν $(\nu \in M+N+L')$. This is what has been done in (18) since the θ's are independent of x^α $(\alpha \in L)$.

From (18), (17), (13) we find that the matrix of the transformation $x \to X$ is

$$P = \left(\frac{\partial x^i}{\partial X^j}\right) = \begin{pmatrix} I & F'-F_1' & G'-G_1' & -K' \\ O & I & O & O \\ O & O & I & O \\ O & O & O & I \end{pmatrix},$$

where primes denote transposes, and the partition here is the same as before.

Writing g for the matrix of the fundamental tensor given by (13), its transform is $P'gP$, and this we find to be

$$\begin{pmatrix} O & O & O & I \\ O & A & O & F \\ O & O & B & G \\ I & F' & G' & C \end{pmatrix}$$

where $C = C_1 - K - K'$. Because of the properties of the $h_{\rho\tau}$ in (17), F is independent of x^α $(\alpha \in L+N)$ and G is independent of x^β $(\beta \in L+M)$. We proved earlier that A is independent of x^α $(\alpha \in L+N)$ and that B is independent of x^β $(\beta \in L+M)$. From the form of the transformation $x \to X$ in (18) we may replace x by X in these statements about independence. Hence, in the coordinate system X^i, the fundamental tensor of V_n satisfies all the conditions of the theorem.

The proof of the theorem is completed by verifying that for the V_n with the fundamental tensor of the theorem, the $(r+s)$-plane p with basis δ^i_ρ $(\rho \in L+M)$ at each point is parallel, and that its null part is the r-plane with basis δ^i_α $(\alpha \in L)$. Sufficient conditions for p to be parallel were found to be (14), equivalent to (16), and these are satisfied. To find the null part of p, we want the vectors of p which are orthogonal to p, and these are easily seen to be the vectors of the r-plane referred to above. This completes the proof of the theorem.

The $(r+s)$-plane of nullity r is *strictly* parallel if a basis can be found consisting of r null and s non-null parallel vector fields, all independent and mutually orthogonal. It follows that V_n is the product of a flat E_s and a V_{n-s}, and that the V_{n-s} admits a strictly parallel null r-plane. The canonical form for such a V_{n-s} was given in C.F.(I) § 7. We finally have for V_n the canonical form

$$(g_{ij}) = \begin{pmatrix} O & O & O & I \\ O & A & O & O \\ O & O & B & G \\ I & O & G' & C \end{pmatrix},$$

where the partition is as before, A is a constant non-singular matrix (the fundamental tensor of E_s) and B, C, G, G' are independent of $x^1,\dots,x^{r+s}$. This form, which is seen to be a special case of our general canonical form, is equivalent to that given by Eisenhart.†

† *Annals of Math.* 39 (1938), 316.

A SIX-VECTOR DEVELOPMENT OF SOME RESULTS IN KINEMATICAL RELATIVITY

By R. A. NEWING (*Bangor*)

[Received 9 April 1949]

MILNE has, in general, explicitly avoided the calculus of space-time as obscuring the physical ideas of his theory. Space-time calculus is, however, a powerful technique which simplifies considerably the mathematical development of many aspects of the theory. In this paper some of Milne's results will be exhibited in 6-vector form. Milne has already introduced 6-vectors in his electromagnetic theory, but it will be shown here that the equations of motion for a free particle, and the equations for gravitational interaction, take on a particularly simple form when expressed as 6-vector equations. The reference 'K.R.' throughout the paper is to Milne's recent book *Kinematical Relativity* (Oxford, 1948).

1. Four-vectors

Milne's fundamental observers have a common space-time defined by the metric

$$ds^2 = dx_1^2+dx_2^2+dx_3^2+dt^2,$$

where $(x_1, x_2, x_3) = (i/c)(x, y, z)$. Let α and β be the 4-vectors $(\mathbf{a}, \alpha_4)$ and $(\mathbf{b}, \beta_4)$, where $\mathbf{a}$ and $\mathbf{b}$ are 3-vectors; then the scalar product $\mathbf{a}.\mathbf{b}+\alpha_4\beta_4$ is an invariant and will be denoted by $\alpha.\beta$. The motion of a particle can be discussed in terms of the 4-vector $R = \left(\frac{i}{c}\mathbf{P}, t\right)$, where $\mathbf{P}$ is its position relative to the observer, and t is the time. The scalar invariants X, $Y^{-\frac{1}{2}}Z$ and ξ may then be expressed as follows:

$$X = t^2-\frac{1}{c^2}|\mathbf{P}|^2 = R.R = R^2,$$

$$Y^{-\frac{1}{2}}Z = \left(1-\frac{V^2}{c^2}\right)^{-\frac{1}{2}}\left(t-\frac{\mathbf{P}.\mathbf{V}}{c^2}\right) = R.R',$$

$$\xi = \frac{Z^2}{XY} = \frac{(R.R')^2}{R^2},$$

where
$$R' = \frac{dR}{ds} = \left(\frac{i\mathbf{V}}{c}, 1\right)\frac{dt}{ds} = \left(\frac{i\mathbf{V}}{c}, 1\right)Y^{-\frac{1}{2}}.$$

It should be noted that $(R'.R') = 1$, and therefore $(R'.R'') = 0$.

2. Six-vectors

A 6-vector (or skew-symmetric tensor of rank 2) Θ may be represented in terms of two 3-vectors $\mathbf{A}$ and $\mathbf{B}$ such that

$$\Theta = [\mathbf{A};\mathbf{B}] = \begin{bmatrix} 0 & A_3 & -A_2 & B_1 \\ -A_3 & 0 & A_1 & B_2 \\ A_2 & -A_1 & 0 & B_3 \\ -B_1 & -B_2 & -B_3 & 0 \end{bmatrix}.$$

The 6-vector $[\alpha_i\beta_j-\alpha_j\beta_i]$, formed from two 4-vectors α and β, will be written as $[\alpha,\beta]$ so that

$$[\alpha,\beta] = [\mathbf{a}\wedge\mathbf{b};\beta_4\mathbf{a}-\alpha_4\mathbf{b}]. \tag{1}$$

The scalar product of two 6-vectors $[\mathbf{A};\mathbf{B}]$ and $[\mathbf{C};\mathbf{D}]$ is an invariant and will be written as $[\mathbf{A};\mathbf{B}]:[\mathbf{C};\mathbf{D}] = (\mathbf{A}.\mathbf{C})+(\mathbf{B}.\mathbf{D})$. In particular

$$\Theta^2 = \Theta:\Theta = |\mathbf{A}|^2+|\mathbf{B}|^2. \tag{2}$$

The product $\Theta\alpha$ is the 4-vector σ, where $\sigma_i = \sum_j \Theta_{ij}\alpha_j$, thus

$$\Theta\alpha = (\mathbf{a}\wedge\mathbf{A}+\alpha_4\mathbf{B}, -\mathbf{a}.\mathbf{B}), \tag{3}$$

and, if γ is the 4-vector $(\mathbf{c},\gamma_4)$,

$$[\alpha,\beta]\gamma = (\mathbf{c}\wedge(\mathbf{a}\wedge\mathbf{b})+\gamma_4(\beta_4\mathbf{a}-\alpha_4\mathbf{b}), -\mathbf{c}.(\beta_4\mathbf{a}-\alpha_4\mathbf{b})). \tag{4}$$

An alternative form is

$$[\alpha,\beta]\gamma = \alpha(\beta.\gamma)-\beta(\alpha.\gamma). \tag{4'}$$

Further results required are

$$[\alpha,\beta]:[\gamma,\delta] = (\alpha.\gamma)(\beta.\delta)-(\alpha.\delta)(\beta.\gamma)$$

and, in particular,

$$[\alpha,\beta]:[\alpha,\beta] = \alpha^2\beta^2-(\alpha.\beta)^2. \tag{5}$$

Further
$$[\delta,[\alpha,\beta]\gamma] = [\delta,\alpha](\beta.\gamma)-[\delta,\beta](\alpha.\gamma), \tag{6}$$

$$\begin{aligned}(\delta.[\alpha,\beta]\gamma) &= (\delta.\alpha)(\beta.\gamma)-(\delta.\beta)(\alpha.\gamma) \\ &= [\alpha,\beta]:[\delta,\gamma]. \end{aligned} \tag{7}$$

Finally, if $S = [R,R']$,

$$S:S = R^2-(R.R')^2 = X(1-\xi) = -\theta_1^2. \tag{8}$$

But, by (1), S may be represented as

$$S = Y^{-\frac{1}{2}}\left[-\frac{1}{c^2}\mathbf{P}\wedge\mathbf{V};\frac{i}{c}(\mathbf{P}-\mathbf{V}t)\right], \tag{9}$$

so that, using (2),

$$-YS^2 = Y\theta_1^2 = \frac{1}{c^2}|\mathbf{P}-\mathbf{V}t|^2-\frac{1}{c^4}|\mathbf{P}\wedge\mathbf{V}|^2. \tag{8'}$$

3. The motion of a free particle

Following Milne's space-time discussion [K.R., 65–6], write

$$R'' = \lambda R + \mu_1 R',$$

where λ and μ_1 are scalar invariants. Now, if S denotes the 6-vector $[R, R']$, $SR' = R-(R.R')R'$, and the expression for R'' may be written in the equivalent form $R'' = \lambda SR' + \mu R'$. Then, since $R'.R''$ and $R'.SR'$ are both zero, $\mu = R'.R'' - \lambda R'.SR' = 0$, and therefore

$$R'' = \lambda SR' \tag{10}$$

where, as in Milne's discussion, λ may be expressed as $G(\xi)/X$. Using (4'), equation (10) gives the equations of motion in 4-vector form [K.R., 67, (23) and (24)].

From (10), $S' = [R, R''] = \lambda[R, SR']$, and therefore

$$S' = -\frac{G(\xi)(R.R')}{X} S. \tag{11}$$

The 3-vector form of the equations of motion [K.R., p. 67, (26)] follows from (11) using (9) and the time-component of (10). Milne's 3-vector equations, in which the terms are not the space parts of 4-vectors, are thus seen to arise from 6-vector equations. [Cf. Milne, *Phil. Mag.* 36 (1945), 138.] Now from (8), $-\theta_1 \theta_1' = S:S' = -(1/X)G(\xi)(R.R')S^2$, and therefore

$$-\frac{G(\xi)(R.R')}{X} = \frac{\theta_1'}{\theta_1}. \tag{12}$$

Equation (11) can now be expressed in the compact 6-vector form

$$\left(\frac{1}{\theta_1} S\right)' = 0. \tag{13}$$

It may be noted that (10) follows from (13) so that the two equations are completely equivalent, for, since $S' = (\theta_1'/\theta_1)S$,

$$(\theta_1'/\theta_1)SR' = S'R' = R(R''.R') - R''(R.R') = -R''(R.R'),$$

and hence (10) follows.

Equation (13) implies that $S = \theta_1 \Theta_0$ where Θ_0 is a constant 6-vector such that $\Theta_0^2 = -1$. Writing

$$\Theta_0 = \left[-\frac{1}{c}\mathbf{l}; \frac{i}{c}\mathbf{f}\right],$$

where $\mathbf{l}$ and $\mathbf{f}$ are constant 3-vectors, and using the representation (9) for S, the vector integrals at once follow

$$\mathbf{P} - \mathbf{V}t = \theta\mathbf{f}, \qquad \mathbf{P} \wedge \mathbf{V} = c\theta\mathbf{l},$$

with $\qquad \theta = Y^{\frac{1}{2}}\theta_1 = Y^{\frac{1}{2}}(\xi-1)^{\frac{1}{2}}X^{\frac{1}{2}} \quad \text{and} \quad |\mathbf{f}|^2 - |\mathbf{l}|^2 = c^2.$

The X-integral follows from (12) and (8), for

$$-2(R.R')(\xi-1)G(\xi) = \{(\xi-1)X\}',$$

and therefore $$X\xi' = -(\xi-1)X'-X'(\xi-1)G(\xi),$$

and hence $$X = X_0 \exp\left\{-\int_{\xi_0}^{\xi} \frac{d\xi}{(\xi-1)\{1+G(\xi)\}}\right\}.$$

4. Dynamical equations

The mass of a particle is defined to be $M = m\xi^{\frac{1}{2}}$, where m is a constant. In the present notation $M = m(R.R')/|R|$, so that

$$M' = \frac{m}{|R|^3}\{(R.R')'R^2-(R.R')^2\}$$
$$= \frac{M}{R.R'}R.\left(R''+\frac{1}{R^2}SR'\right).$$

If $G(\xi)$ is put equal to -1, M is constant and (10) may be written as

$$(MR')'+\frac{M}{R^2}SR' = 0. \tag{14}$$

If (14) is not satisfied, the 4-vector

$$\frac{1}{c}(i\mathbf{F}, F_t) = Q = (MR')'+\frac{M}{R^2}SR' \tag{15}$$

defines the force acting on the particle. Then

$$Q.R' = M', \tag{16}$$

giving at once the relation between F_t and $\mathbf{F}$ [K.R., 82, (5)]

$$cF_t-\mathbf{F}.\mathbf{V} = c^2Y^{\frac{1}{2}}M' = \frac{d}{dt}(mc^2\xi^{\frac{1}{2}}).$$

It is to be noted that $Q = 0$ implies $M' = 0$, so that (14) is completely equivalent to (10) with $G(\xi) \equiv -1$.

Equation (15) also implies that

$$Q.R = 2(R.R')M', \tag{17}$$

and therefore $$Q.\left(\frac{1}{R.R'}R-R'\right) = M'.$$

The last equation is

$$\mathbf{F}.(\mathbf{V}Y^{-\frac{1}{2}}-\mathbf{P}Y^{\frac{1}{2}}Z^{-1})-cF_t(Y^{-\frac{1}{2}}-tY^{\frac{1}{2}}Z^{-1}) = Y^{-\frac{1}{2}}\frac{d}{dt}(mc^2\xi^{\frac{1}{2}}),$$

and has been interpreted as an energy equation [K.R., 84], the left-hand side giving the rate at which work is done in moving the

particle relative to its surroundings in the substratum. Equation (15) may be written in the equivalent form

$$[R, Q] = |R|\left(\frac{M}{|R|}S\right)',$$

which gives, when $\mathbf{P}\wedge\mathbf{F} = 0$, the angular-momentum integral [K.R., 94, (52)] in the form $M\mathbf{P}\wedge\mathbf{V} = \mathbf{A}|R|Y^{\frac{1}{2}}$, where $\mathbf{A}$ is a constant 3-vector.

Consider now the force defined by the 4-vector

$$Q = \Theta R + \Phi R' + aR' + bR,$$

where Θ, Φ are 6-vectors and a, b are scalar invariants. Without loss of generality we may take $b = 0$ since $R = SR' + R'(R.R')$. Then, if

$$Q = \Theta R + \Phi R' + aR',$$

(16) implies that
$$a = M' - R'.\Theta R, \tag{18}$$

and, from (17),
$$a = 2M' - \frac{1}{R.R'}R.\Phi R'. \tag{19}$$

Equation (15) may now be written in the 6-vector form

$$MS' - \left\{M' + \frac{M(R.R')}{R^2}\right\}S = [R, \Theta R] + [R, \Phi R'] - \frac{R.\Phi R'}{R.R'}S,$$

which reduces to

$$M(R.R')\left(\frac{1}{R.R'}S\right)' = [R, \Theta R] - \frac{1}{R.R'}[R, [R', \Phi R']R]. \tag{20}$$

Two special cases are important for the discussion of gravitational and electromagnetic interactions between particles:

Case (i)

$$Q = \Theta R + aR', \qquad a = 2M' = -2R'.\Theta R,$$

$$M(R.R')\left(\frac{1}{R.R'}S\right)' = [R, \Theta R]. \tag{21}$$

Case (ii)

$$Q = \Phi R' + aR', \qquad a = M' = \frac{1}{R.R'}(R.\Phi R'),$$

$$M(R.R')\left(\frac{1}{R.R'}S\right)' + M'S = [R, \Phi R']. \tag{22}$$

The latter equation may be written more simply as

$$M|R|\left(\frac{1}{|R|}S\right)' = [R, \Phi R']. \tag{22'}$$

5. Gravitational interaction between particles

Consider two particles defined by the mass-constants m_1, m_2 and the 4-vectors R_1, R_2. Let S_1, S_2, and S_{12} denote the 6-vectors $[R_1, R_1']$, $[R_2, R_2']$ and $[R_1, R_2]$ respectively. The simplest 6-vector equation for the motion of m_1 in which the interaction does not involve the velocity of m_2 is

$$M_1(R_1 . R_1')\left(\frac{1}{R_1 . R_1'} S_1\right)' = \lambda_{12}[R_1, R_2], \tag{23}$$

λ_{12} being a scalar invariant symmetrical in the suffixes 1 and 2. Then, since the force is of type (i), $[R_1, \Theta_1 R_1] = \lambda_{12} S_{12}$ and therefore

$$\Theta_1 R_1 = -(\lambda_{12}/R_1^2) S_{12} R_1.$$

There will be similar equations for m_2, and the forces on the two particles are thus defined by the 4-vectors $-\mu R_2^2 S_{12} R_1$ and $-\mu R_1^2 S_{21} R_2$, where λ_{12} is put equal to $R_1^2 R_2^2 \mu$. For forces of type (i), $M' = -R'.\Theta R$, and therefore

$$\frac{d}{dt}(M_1+M_2) = \mu\left\{R_2^2 \frac{dR_1}{dt}.S_{12} R_1 + R_1^2 \frac{dR_2}{dt}.S_{21} R_2\right\},$$

where t is a parameter dependent upon the simultaneity convention adopted. Now, as in Milne's treatment, this last equation may be made to give the energy integral

$$M_1 c^2 + M_2 c^2 + \chi = \text{constant}$$

by taking μ to be $m_1 m_2/M_0 \alpha_{12}^3$, with

$$\alpha_{12}^2 = X_{12}^2 - X_1 X_2 = (R_1 . R_2)^2 - R_1^2 R_2^2 = -S_{12}^2.$$

χ is then taken as $-(m_1 m_2 c^2/M_0 \alpha_{12})(R_1 . R_2)$, and this gives

$$\frac{\partial \chi}{\partial R_1} = c^2 \Theta_1 R_1, \qquad \frac{\partial \chi}{\partial R_2} = c^2 \Theta_2 R_2.$$

For a system of particles (23) becomes

$$M_1(R_1 . R_1')\left(\frac{1}{R_1 . R_1'} S_1\right)' = \sum_{r=1}^{n} \lambda_{1r} S_{1r},$$

which, since $\lambda_{rs} = \lambda_{sr}$, $S_{rs} = -S_{sr}$, gives when summed over all the particles of the system

$$\sum_{r=1}^{n} M_r(R_r . R_r')\left(\frac{1}{R_r . R_r'} S_r\right)' = 0.$$

If the 6-vectors S_r are expressed as in (9), this last equation at once gives Milne's momentum equations

$$\sum_{r=1}^{n} M_r(R_r.R'_r)Y_r^{-\frac{1}{2}}\frac{d}{dt_r}\left(\frac{\mathbf{V}_r t_r-\mathbf{P}_r}{Z_r}\right) = 0,$$

$$\sum_{r=1}^{n} M_r(R_r.R'_r)Y_r^{-\frac{1}{2}}\frac{d}{dt_r}\left(\frac{\mathbf{P}_r\wedge\mathbf{V}_r}{Z_r}\right) = 0.$$

6. Electromagnetic interaction between particles

The representation of the force on a charge e_1 due to a charge e_2 in terms of the 6-vector $[\mathbf{H}_1; i\mathbf{E}_1]$ is well known, and it is sufficient to note that, in the present notation, this 6-vector takes the form

$$[\mathbf{H}_1; i\mathbf{E}_1] = \frac{-e_2}{2c^2\alpha_{12}^3}[R'_2(R_2.R'_1)+R'_1(R_2.R'_2),[R_1,R_2]R_2],$$

for equation (1) at once gives Milne's expressions [K.R., 180, (37), (38)] for $\mathbf{H}_1$ and $\mathbf{E}_1$.

If the gravitational interaction is neglected, the force on e_1 is of type (ii), so that $a_1 = M'_1$ and

$$\Phi_1 = \frac{e_1(R_1.R'_1)}{ct_0}[\mathbf{H}_1; i\mathbf{E}_1].$$

For combined gravitational and electromagnetic interactions

$$Q_1 = \Theta_1 R_1+\Phi_1 R'_1+a_1 R'_1,$$

and, from (18) and (19),

$$M'_1+R'_1.\Theta_1 R_1 = \frac{1}{R_1.R'_1}R_1.\Phi_1 R'_1 = \frac{1}{R_1.R'_1}\Phi_1\!:[R_1,R'_1].$$

This gives at once the energy equation

$$c^2M'_1+\frac{\partial\chi}{\partial s_1} = \frac{e_1}{t_0}\left\{\mathbf{E}_1.(\mathbf{V}_1 t_1-\mathbf{P}_1)-\frac{1}{c}\mathbf{H}_1.(\mathbf{P}_1\wedge\mathbf{V}_1)\right\}Y_1^{-\frac{1}{2}}$$

$$= \frac{e_1 t_1}{t_0 Y_1^{\frac{1}{2}}}\left(\mathbf{E}_1+\frac{1}{c}\mathbf{V}_1\wedge\mathbf{H}_1\right).\left(\mathbf{V}_1-\frac{1}{t_1}\mathbf{P}_1\right).$$

The 6-vector $[R_1,\Phi_1 R'_1]$ can be expressed in terms of S_1, S_{12}, and $[R_1, R'_2]$, and the right-hand sides of (20) and (22′) take the form $\lambda_1 S_1+\mu_1 S_{12}+\nu_1[R_1, R'_2]$, where λ_1, μ_1, and ν_1 are scalar invariants. The electromagnetic equations, however, have not the simplicity of the corresponding equations for gravitational interaction. If the particle e_2 is much more massive than e_1, we may take e_2 coincident with a fundamental particle and put $R_2 = (0, t_2)$, $R'_2 = (0, 1)$. The 6-vectors

S_{12} and $[R_1, R_2']$ are then both of the form $[\mathbf{O}; \mathbf{B}]$, and (22′) reduces to the form

$$M_1|R_1|\left(\frac{1}{|R_1|}S_1\right)' = \frac{e_1 e_2(R_1 . R_1')(\mathbf{P}_1 . \mathbf{V}_1)}{2t_0|\mathbf{P}_1|^3 Y_1^{\frac{3}{2}} c^2} S_1 + \kappa\left[\mathbf{O}; \frac{i}{c}\mathbf{P}_1\right].$$

Expressing S_1 as in (9), it follows that

$$\frac{d}{dt_1}\left(\frac{\mathbf{P}_1 \wedge \mathbf{V}_1}{Y_1^{\frac{1}{2}}|R_1|}\right) = \frac{e_1 e_2}{2m_1 c^2}\frac{\mathbf{P}_1 \wedge \mathbf{V}_1}{Y_1^{\frac{3}{2}} t_0}\frac{(\mathbf{P}_1 . \mathbf{V}_1)}{|\mathbf{P}_1|^3} = \frac{-e_1 e_2}{2m_1 c^2}\frac{\mathbf{P}_1 \wedge \mathbf{V}_1}{Y_1^{\frac{1}{2}}|R_1|}\frac{|R_1|}{t_0}\frac{d}{dt_1}\left(\frac{1}{|\mathbf{P}_1|}\right),$$

which gives at once Milne's angular momentum integral [K.R., 205, (22)] for the Kepler problem associated with purely electromagnetic interactions.

A THEOREM ON INFINITE PRODUCTS OF TRANSFINITE CARDINAL NUMBERS

(CORRECTION)*

By F. BAGEMIHL (*Rochester, N.Y.*)

[Received 15 October 1949]

Formula (10) on page 202 does not necessarily follow from (4), because $\{\sigma_\xi\}_{\xi<\beta}$ may not satisfy the same conditions that $\{\sigma_\xi\}_{\xi<\alpha}$ does. If, however, the theorem is modified by imposing on the sequence $\{\sigma_\xi\}_{\xi<\alpha}$ the following additional condition, the proof is correct:

If, for $n > 2$,

$$\alpha_r = \sum_{1<k<r} \omega^{\delta_k}, \qquad \beta_r = \sum_{1<k<r-1} \omega^{\delta_k}, \qquad \lim_{\xi<\alpha_r} \sigma_\xi = \lambda_r, \qquad \lim_{\xi<\beta_r} \sigma_\xi = \eta_r,$$

then $-\eta_r + \lambda_r \leqslant \omega^{\delta_r}$ $(r = 2, 3, \ldots, n-1)$.

The error was pointed out to me by P. Erdős.

* See this Journal, 19 (1948), 200–3.